# Enumeration of tanglegrams

by

Jean Bernoulli Ravelomanana

*Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Mathematics in the Faculty of Science at Stellenbosch University*

Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Supervisors:

Dr. Dimbinaina Ralaivaosaona
Prof. Stephan Wagner

March 2018

# Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: ....................................
Jean Bernoulli Ravelomanana

Date: March 2018 ........................................

# Abstract

**Enumeration of tanglegrams**

Jean Bernoulli Ravelomanana

*Department of Mathematical Sciences,*
*University of Stellenbosch,*
*Private Bag X1, Matieland 7602, South Africa.*

Thesis: MSc

December 2017

Tanglegrams are graphs obtained by taking two binary rooted trees with the same number of leaves and a perfect matching between the leaves of the two trees. Tanglegrams appear in biology in the study of cospeciation or coevolution, and in computer science in the study of software projects and clustering problems. This thesis is concerned with the enumeration of tanglegrams: we first prove an exact formula for the number of non-isomorphic tanglegrams on $n$ leaves and an asymptotic formula for the same quantity as $n$ tends to infinity. Next, we study several parameters of random tanglegrams such as the number of occurrences of subtrees or the distribution of root branches. Finally, our main contribution in this thesis is on the enumeration of planar tanglegrams on $n$ leaves, where a planar tanglegram is a tanglegram that can be drawn in the plane without crossings.

# Uittreksel

**Eienskappe van die gulsige bome**

Jean Bernoulli Ravelomanana

*Departement Wiskundige Wetenskappe,*
*Universiteit van Stellenbosch,*
*Privaatsak X1, Matieland 7602, Suid Afrika.*

Tesis: MSc

Desember 2017

Tanglegramme is grafieke wat uit twee binêre wortelbome met dieselfde aantal blare en 'n perfekte matching tussen die blare van die twee bome bestaan. Tanglegramme verskyn in biologie in die studie van kospesiasie en koëvolusie, en in rekenaarwetenskap in die ondersoek van sagtewareprojekte en groeperingsprobleme. Hierdie proefskrif behandel die aftelling van tanglegramme: ons bewys eers 'n formule vir die aantal van nie-isomorfe tanglegramme met $n$ blare en 'n asimptotiese formule vir hierdie aantal as $n$ na oneindig strewe. Verder bestudeer on 'n verskeidenheid van parameters van lukrake tanglegramme soos die aantal voorkoms van deelbome of die verdeling van worteltakke. Laastens is ons hoofbydrag in hierdie proefskrif die aftelling van planêre tanglegramme met $n$ blare, waar 'n planêre tanglegram 'n tanglegram is wat in die vlak sonder kruisings geteken kan word.

# Acknowledgements

I am very grateful to my two supervisors Dr. Dimbinaina Ralaivaosaona and Prof. Stephan Wagner for their continuous support and their advice throughout the entire research period. My sincere thanks go to the Faculty of Science of Stellenbosch University and the African Institute for Mathematical Sciences (AIMS) for financially and materially supporting this project. I am grateful to all my friends , especially Nattie for proof readings, my officemates Ken, Taboka and Valisoa for helpful discussions. My appreciation also goes to my dear family: my two sisters Joliot and Riccati, and my brother Huygens.

# Dedications

*To my big brother and in memory of my parents,*

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction and preliminary results

The concept of pairs of phylogenetic (meaning leaf-labeled) trees with a relative mapping between the sets of leaves has been introduced as tanglegrams in [25] and [26]. Formally, we define a tanglegram as a pair of binary rooted trees with the same number of leaves and a bijection between the sets of leaves. Here, the bijection is represented by inter-tree edges. Tanglegrams appear naturally in biology, in the study of cospeciation and coevolution. For instance, one tree may correspond to the phylogeny of a host, such as a mouse, and the other tree may correspond to a parasite such as a louse, see [2, 29, 34] for more details.

Tanglegrams also appear in computer science. More precisely, they play important roles in the analysis of software projects and clustering problems. For both computer science and biology, an important question is the *Tanglegram Layout* (*TL*) problem which is to find a drawing of a tanglegram where the two trees are both given as planar embeddings with the minimum number of crossings between inter-tree edges. The *TL* problem is important for visualization purposes. For example, in biology, the goal is to see as clearly as possible the coevolutionary relationship between species.

Even though there is a considerable work on the *TL* problem, see for example [1, 4, 7, 12, 33], there has not been much work done on enumerating or finding other properties tanglegrams until recently as it was pointed out in [23]. This latter fact motivates our work which is on the enumeration of tanglegrams. This thesis is organized as follows: in this chapter, we formalize general concepts of tanglegrams, describe their properties and observe that tanglegrams have a useful formulation as double cosets of the symmetric group.

In the second chapter, we use the correspondence between double cosets of the symmetric group and tanglegrams to derive an exact formula for the number of tanglegrams with $n$ leaves. This formula was established by Billey, Konvalinka and Matsen in [3] alongside with an alternative formula for the number of rooted binary trees with $n$ leaves. Furthermore, we extend the concept of tanglegrams to more than two trees and obtain new combinatorial objects called *tangled chains*. Using the approach for tanglegrams, we derive an exact formula for the number of tangled chains. We remark here that many essential problems in phylogenetics can be cut down to questions on labeled sets of more than two trees which, in turn, will correspond to tangled chains, see [34]. In addition, we give an asymptotic formula for

the number of tanglegrams with exactly $n$ leaves in each tree.
In chapter three, following the work of Konvalinka and Wagner in [21], we will see that a random tanglegram looks like two independently chosen random plane binary trees. This fact will be used as a basis to determine the behavior of various parameters of a tanglegram such as the number of occurrences of subtrees, the distribution of root branches, the number of automorphisms and the height. It was said in [3] that cherries (a subtree of a binary tree $T$ consisting of an internal vertex with exactly two leaves as children) play a major role in the literature of tanglegrams. The average number and the limiting distribution of matched cherries (two cherries whose leaves are matched to each other) will be investigated.
In chapter four, we consider the TL problem from the enumerative point of view. More precisely, we aim to answer the question: how many tanglegrams can be drawn without crossings? We call these tanglegrams *planar tanglegrams*. We discover several new results: first, a bijection between a special class of planar tanglegrams and pairs of triangulations of polygons without common diagonals is established. Second, using the previous bijection, we obtain different functional equations for the generating functions of planar tanglegrams. Finally, we use singularity analysis to determine the asymptotic number of planar tanglegrams.
Now, we begin with a review of basic concepts in graph theory and group theory that will be useful in the study of tanglegrams. Those notions can be found in standard graph theory textbooks like [6] or [8] and group theory textbooks such as [11].

## 1.1 Automorphism of rooted trees and tanglegrams

Recall that a tree is a connected simple graph with no cycles. We say that the tree is *rooted* if we distinguish one particular vertex from all other vertices; we call this vertex the *root*. The vertices with degree one are called *leaves* and all other vertices are called *internal vertices*. The vertices adjacent to the root are called children or successors of the root; the vertices that are adjacent to the children of the root (which are not the root) are their children and we continue recursively. A branch of a rooted tree is then the subtree induced by one child of the root and all its successors (if they exist).
Moreover, a binary tree is a rooted tree where every internal vertex has two children (see Figure 1.1). We note that there is no specified order between the right and left child of a vertex in our binary trees. From the previous definition of binary trees, we formalize the concept of tanglegram as it was given in [3] and [23].

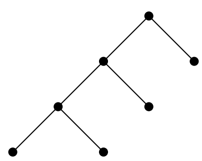

Figure 1.1: A binary tree.

**Definition 1.1.1.** A tanglegram is a pair of binary trees $T$, $S$ with the same number of leaves and a bijection $\phi$ between the leaves of $T$ and $S$. The tanglegram is ordered if the order in which appear $T$ and $S$ matters; in that case, the tanglegram is denoted by the triplet $(T, \phi, S)$. Otherwise, in the unordered case, the tanglegram is denoted by $(\{T, S\}, \phi)$.

Here, tanglegrams are ordered unless it is specified otherwise. Also a tanglegram is drawn with one tree on top and the other tree on the bottom. The bijection $\phi$ is represented by inter-tree edges (see Figure 1.2) and the size of a tanglegram is the number of leaves in each tree.

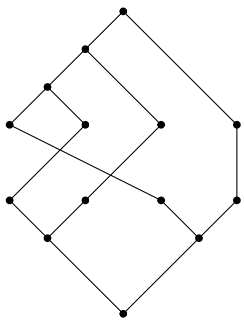

Figure 1.2: A tanglegram of size 4.

Next, let $G = (V, E)$ and $G' = (V', E')$ be two graphs. An isomorphism between $G$ and $G'$ is a bijection $f : V \to V'$ which preserves adjacency i.e $\{a, b\} \in E$ if and only if $\{f(a), f(b)\} \in E'$. In that case, we say that the two graphs $G = (V, E)$ and $G' = (V', E')$ are isomorphic. If $G = G'$, we say that $f$ is an automorphism. For example, the trees in Figure 1.3 are isomorphic. It is clear that for a given pair of isomorphic binary trees $T$ and $S$, the root of $T$ is mapped to the root of $S$ and a leaf of $T$ is also mapped to a leaf of $S$. Furthermore, we have the property that an automorphism of a tree is determined by the bijection between the leaves as it is stated in the next proposition.

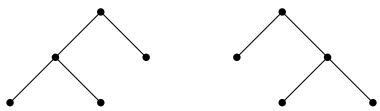

Figure 1.3: Two isomorphic binary trees.

**Proposition 1.1.2** ([23])**.** An isomorphism $f$ between two trees $T$ and $S$ is uniquely determined by the bijection between the set of leaves of $T$ and $S$. In particular, if $g$ is an automorphism then $g$ is determined by the bijection on the set of leaves.

*Proof.* Let $T$ and $S$ be two isomorphic trees, $f$ and $g$ be two isomorphisms that induce the same bijection on the set of leaves. Consider an internal vertex $a$ of $T$. The vertex $a$ lies on a path $P$ between two leaves $x$ and $y$. Indeed, if we take two arbitrary leaves in the two branches of $a$, then the path $P$ between $x$ and $y$ contains a. Since isomorphisms preserve adjacency, they send a path to a path. Thus, $f$ and $g$ map the path $P$ to paths $P_1$ and $P_2$ between the leaves $f(x)$ and $f(y)$ ($g(x)$ and $g(y)$ respectively) . Since $f(x) = g(x)$ and $f(y) = g(y)$, the two paths $P_1$ and $P_2$ must be the same by definition of a tree. Hence $f(a) = g(a)$. □

From now on, we consider isomorphisms between trees as bijections between the sets of leaves. We remark that the set of all automorphisms of a graph $G = (V, E)$ forms a group under composition denoted by $A(G)$. In particular, if a tree $T$ has $n$ leaves then the automorphism group $A(T)$ is a subgroup of the symmetric group $S_n$. In order to understand trees and thus tanglegrams, we study the structure of these automorphism groups. To this end, we need to define the so called *wreath product* of two groups.
Let $G$ be a group and $H$ be a permutation group on a set $X$. Let $G^X$ be the set of functions from $X$ to $G$ which we equip with the pointwise operation

$$(f \cdot g)(x) = f(x)g(x),$$

for $f, g \in G^X$ and $x \in X$. With these definitions, it is easy to check that $(G^X, \cdot)$ forms a group. Now, consider the set $C = G^X \times H$. We define an operation on $C$.

**Definition 1.1.3.** For $(f, h), (f', h') \in C$,

$$(f, h) \star (f', h') = (f^{(h')} \cdot f', hh'),$$

where $f^{(h')}(x) = f(h'(x))$ for $x \in X$.

Note that for $(f, h), (f', h') \in C$ and $x \in X$, we have:

- $(f \cdot f')^{(h)}(x) = (f \cdot f')(h(x)) = f(h(x))f'(h(x)) = f^{(h)}(x) \cdot f^{(h')}(x)$ and
- $f^{(hh')}(x) = (f^{(h)})^{(h')}(x) = f^{(h)}((h'(x))) = f(h(h'(x)))$.

We have the following proposition:

**Proposition 1.1.4** ([16], p. 81)**.** The set $C$ together with the operation $\star$ form a group called *wreath product* of $G$ by $H$, denoted by $G \wr H$.

*Proof.* Here we use the same symbol " $\cdot$ " for the operation of $G^X$ , $G$ and $H$. The associativity of $\star$ comes from the associativity of the operations of $G^X$ and $H$. Let $I$ be the identity element in $G^X$, i.e. $I(x) = e$ for all $x \in X$, where $e$ is the neutral element of $G$, let $Id$ be the identity element of $H$ and $(f, h) \in C$. Then, $(I, Id) \star (f, h) = (I^{(h)} \cdot f, Id \cdot h) = (I^{(h)} \cdot f, h)$. We have $(I^{(h)} \cdot f)(x) = I(h(x)) \cdot f(x) = e \cdot f(x) = f(x)$. So, $(I, Id) \star (f, h) = (f, h)$. Similarly, $(f, h) \star (I, Id) = (f^{(Id)} \cdot I, h \cdot Id) = (f^{(Id)} \cdot I, h)$. We have $(f^{(Id)} \cdot I)(x) = f(Id(x)) \cdot I(x) = f(x) \cdot e = f(x)$. Thus, $(f, h) \star (I, Id) = (f, h)$ and $(I, Id)$ is the neutral element of $C$.
Now, for $(f, h) \in C$ define $f_h^{(-1)}$ by

$$f_h^{(-1)}(x) = (f^{(h^{-1})}(x))^{-1} = (f(h^{-1}(x)))^{-1}.$$

Then,

$$(f_h^{(-1)}, h^{-1}) \star (f, h) = ((f_h^{(-1)})^{(h)} \cdot f, h^{-1} \cdot h) = (I, Id).$$

Indeed,

$$\left((f_h^{(-1)})^{(h)} \cdot f\right)(x) = (f(h^{-1}(h(x))))^{-1} \cdot f(x) = (f(x))^{-1} \cdot f(x) = e = I(x).$$

In the same manner, we also have

$$(f, h) \star (f_h^{(-1)}, h^{-1}) = (I, Id).$$

Hence $(f_h^{(-1)}, h^{-1})$ is the inverse of $(f, h)$. □

For our purposes, the set $X$ will be a finite set of cardinality say $k$. The group $G^X$ is then identified with the *k-fold direct product* of $G$ denoted by $G^k$ with component-wise operation and the group $H$ is the symmetric group $S_k$ of all permutations of $X$. Given elements $g, g'$ in $G^k$ and $\sigma, \sigma'$ in $S_k$, as it was stated earlier, the operation on $G^k \wr S_k$ is given by

$$(g, \sigma) \star (g', \sigma') = (g^{(\sigma')} g', \sigma\sigma').$$

**Remark 1.1.5.** When we identify $G^X$ with $G^k$, the operation $g^{(\sigma')}$ behaves in the following way: each component $g_i$ of $g$ is permuted to the component $g_{\sigma'(i)}$ for $g = (g_1, g_2, \ldots, g_k) \in G^k$ and $\sigma' \in S_k$. Indeed, suppose $X = \{x_1, x_2, \ldots, x_k\}$ (ordered following the indices $i \in \{1, \ldots, k\}$) and let $g : X \to G$ be a function. Then, $g$ is identified with the $k$-uple $(g(x_1), g(x_2), \ldots, g(x_k))$. Hence, applying a permutation $\sigma'$ to $g$ exchanges $g(x_i)$ and $g(\sigma'(x_i))$, i.e. $g^{(\sigma')}$ corresponds to $\Big(g(\sigma'(x_1), g(\sigma'(x_2)), \ldots, g(\sigma'(x_k))\Big)$.

Wreath products of the form $G \wr S_k$ characterize the automorphism groups of rooted trees as it is stated in the next theorem due to Jordan (see [20] and [23]).
Let $T$ be a rooted tree where the root has $k$ children. Let $T_1, \ldots, T_k$ be the $k$ branches. We rearrange those $k$ branches, with respect to isomorphism, into a partition:

$$P = \{\{T_1, \cdots, T_{i_1}\}, \{T_{i_1+1}, \cdots, T_{i_2}\}, \cdots, \{T_{i_{p-1}+1}, \cdots, T_{i_p}\}\}, \tag{1.1.1}$$

where each part is composed of isomorphic trees, $i_j - i_{j-1}$ is the number of isomorphic trees in the part containing $T_{i_j}$ (assuming $i_0 = 0$) and $p$ is the total number of parts.

**Theorem 1.1.6** (Jordan, 1869)**.** *The automorphism group $A(T)$ of $T$ is given by the direct product $A_{i_1} \times A_{i_2} \times \cdots \times A_{i_p}$, where $A_{i_j}$ is the wreath product $A(T_{i_j}) \wr S_{i_j - i_{j-1}}$.*

*Proof.* Consider a part $P_{i_j} = \{T_{i_{j-1}+1}, \ldots, T_{i_j}\}$ in $P$. All the subtrees in $P_{ij}$ are isomorphic, which means that all the automorphism groups of the subtrees in $P_{i_j}$ are the same as the automorphism group of a given subtree $T_k \in P_{i_j}$. This fact corresponds to a direct product $A(T_k)^{i_j - i_{j-1}}$. Moreover, each subtree $T_s \in P_{i_j}$ can be mapped to itself or to another subtree $T'_s$ which corresponds to a permutation of two components of an element $g \in A(T_k)^{i_j - i_{j-1}}$ by a permutation $\sigma \in S_{i_j - i_{j-1}}$. That is, the symmetry group corresponding to $P_{i_j}$ is $i_j - i_{j-1}$ copies of $A(T_k)$

along with the symmetric group that permutes the isomorphic trees in $P_{i_j}$, equipped with the group operation that appropriately exchanges the subtrees before applying isomorphisms to the subtrees. This symmetry group is the wreath product $A(T_{i_j}) \wr S_{i_j - i_{j-1}}$ (see Remark 1.1.5). Since elements of different parts are not isomorphic, then $A(T) = A_{i_1} \times A_{i_2} \times \cdots \times A_{i_p}$. □

**Corollary 1.1.7** ([23])**.** The automorphism group of a binary tree can be obtained by iterated direct and wreath products of $\mathbb{Z}_2$.

*Proof.* Let $T$ be a binary tree, $T_1$ and $T_2$ be the corresponding branches. From Theorem 1.1.6, if $T_1$ is not isomorphic to $T_2$, then $A(T) = A(T_1) \times A(T_2)$, otherwise $A(T) = A(T_1) \wr \mathbb{Z}_2$. So, we can recursively construct the automorphism group of a binary tree starting from the tree with one vertex (which has trivial automorphism group) and taking direct and wreath products of $\mathbb{Z}_2$. □

**Example 1.1.8.** In Figure 1.3, in the left tree, the automorphism group corresponding to the left branch is $\mathbb{Z}_2$ and the automorphism group corresponding to the right branch is the trivial group $\{1\}$. Thus, the automorphism group of the entire tree is $\{1\} \times \mathbb{Z}_2 \cong \mathbb{Z}_2$. A similar reasoning applies to the right tree so that the corresponding automorphism group is $\mathbb{Z}_2$ which agrees with the fact that the two trees are isomorphic.

## 1.2 Double cosets

Proposition 1.1.2 provides us with a way to identify tanglegrams that are equivalent and motivates the following definition which is given in [3] and [23]. We denote by $L(T)$ the set of leaves of a tree $T$ and by $|T|$ the number of leaves.

**Definition 1.2.1.** Two tanglegrams $X = (T, \phi, S)$ and $X' = (T, \phi', S)$ on the same set of trees are isomorphic if there exist two automorphisms $f : L(T) \to L(T)$ and $g : L(S) \to L(S)$ such that $g \circ \phi = \phi' \circ f$. In other words, the following diagram is commutative:

$$\begin{tikzcd} L(T) \arrow[r, "\phi"] \arrow[d, "f"] & L(S) \arrow[d, "g"] \\ L(T) \arrow[r, "\phi'"] & L(S). \end{tikzcd}$$

**Remark 1.2.2.** From the previous definition, we have $\phi = g^{-1} \circ \phi' \circ f$. Now, we use the same set of labels on the leaves of $T$ and $S$ such that $A(T)$ and $A(S)$ can be identified as subgroups of $S_n$ ($|T| = |S| = n$). Then, the previous definition says that $\phi$ is an element of the set $\{h \circ \phi' \circ k | h \in A(T) \text{ and } k \in A(S)\}$; such sets are called double cosets ([19]) of $S_n$ with respect to $A(S)$, $A(T)$ and $\phi'$.

Let $G$ be a group and $H, K$ be subgroups of $G$.

**Definition 1.2.3.** For $g \in G$, the set $HgK = \{jgk | j \in H \text{ and } k \in K\}$ is called a double coset of $G$, with respect to $H$ and $K$.

We have the following proposition which comes from Remark 1.2.2 (see also [23]).

**Proposition 1.2.4** ([23])**.** Let $T$ and $S$ be two binary trees with $n$ leaves.

- The set of tanglegrams isomorphic to a tanglegram $(T, \phi, S)$ is in one-to-one correspondence with the double coset $A(S)\phi A(T)$.
- The set of unordered tanglegrams isomorphic to an unordered tanglegram $(\{T, S\}, \phi)$ is in one-to-one correspondence with the equivalence class of the double coset $A(S)\phi A(T)$ where the double cosets $A(S)\phi A(T)$ and $A(T)\phi^{-1}A(S)$ are considered equivalent.

**Remark 1.2.5.** Let $x, y \in G$. The relation defined by $x \sim y$ if and only if $y = hxk$ for some $h \in H$ and $k \in K$ is an equivalence relation. The equivalence class of an element $x \in G$ is the double coset $HxK$. Equivalently, the set of double cosets with respect to $H$ and $K$ partitions the group $G$.

Recall the following well known result in group theory relating the cardinality of $HK$ with the cardinality of $H$ and $K$ if they are finite.

**Proposition 1.2.6** ([19])**.** If $H$ and $K$ are finite then we have:

$$|HK| = \frac{|H||K|}{|H \cap K|} = |H| \cdot [K : H \cap K]. \tag{1.2.1}$$

This leads to the corollary:

**Corollary 1.2.7** ([19])**.** Suppose $H$ and $K$ are finite. Then, for $x \in G$ we have:

$$|HxK| = \frac{|H||K|}{|H \cap xKx^{-1}|} = |H| \cdot [K : H \cap xKx^{-1}]. \tag{1.2.2}$$

*Proof.* The map $F : HxK \to HxKx^{-1}$ defined by $F(hxk) = hxkx^{-1}$ is a bijection. Moreover, since $xKx^{-1}$ is a finite group, from Equation (1.2.1) we have:

$$|HxK| = |HxKx^{-1}| = \frac{|H||xKx^{-1}|}{|H \cap xKx^{-1}|}.$$

Since the map $F' : K \to xKx^{-1}$ defined by $F'(k) = xkx^{-1}$ is also a bijection, the previous equation gives Equation (1.2.2). □

The previous corollary is already quite useful for counting non-isomorphic tanglegrams as it is stated in the following proposition found in [23].

**Proposition 1.2.8** ([23])**.** For two binary trees $T$ and $S$ with the same number of leaves $n$, the number of tanglegrams isomorphic to a tanglegram $X = (T, \phi, S)$ is equal to

$$|A(S)|[A(T) : A(T) \cap \phi^{-1}A(S)\phi] = |A(T)|[A(S) : A(S) \cap \phi A(T)\phi^{-1}],$$

or equivalently,

$$|A(T)\phi A(S)| = \frac{|A(T)| \cdot |A(S)|}{|A(T) \cap \phi A(S)\phi^{-1}|}.$$

*Proof.* From Proposition 1.2.4, the set of tanglegrams isomorphic to a tanglegram $(T, \phi, S)$ is in one-to-one correspondence with the double coset $A(S)\phi A(T)$. Thus, Proposition 1.2.8 is obtained by applying Corollary 1.2.7 to $A(S)\phi A(T)$. $\square$

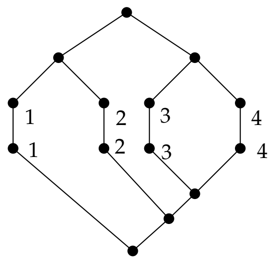


Figure 1.4: A labelled tanglegram with 4 leaves.

**Example 1.2.9.** Consider the tanglegram in Figure 1.4. Let $T$ be the top tree and $S$ be the bottom tree. We put the same set of labels $\{1, 2, 3, 4\}$ on the leaves of the two trees. The automorphism groups of the two branches of $T$ are both $\mathbb{Z}_2$. Since the two branches are isomorphic, the automorphism group of $T$ is $A(T) = (\mathbb{Z}_2 \times \mathbb{Z}_2) \wr \mathbb{Z}_2$. The automorphism group of $A(T)$ can also be viewed as the subgroup of $S_4$ generated by $V = \{(1, 2), (3, 4), (1, 3)(2, 4)\}$. The first two permutations in $V$ interchange the leaves of a branch, while the third interchanges the two branches. For the tree $S$ we can only exchange the leaves 3 and 4 so the automorphism group of $A(S)$ is $\{(), (3, 4)\} \cong \mathbb{Z}_2$ ("()" is the identity permutation). We have $|A(T)| = 8$, so by Proposition 1.2.8, for any permutation $\phi \in S_4$, 8 divides $|A(T)\phi A(S)|$. Since $|S_4| = 24$ and the set of double cosets with respect to $A(T)$ and $A(S)$ partitions $S_4$ (Remark 1.2.5) we only have three possible cases: three double cosets of cardinality 8, two double cosets (one of cardinality 8 and one of cardinality 16) and one double cosets of cardinality 24. If we let $\phi = (2, 3)$ then, using Sagemath ([32]), we have:

$$A(T)\phi A(S) = \{(2,3), (2,4,3), (1,2,4,3), (1,2,3), (2,3,4), (2,4), (1,3,2),$$
$$(1,4,3,2), (1,3,4,2), (1,4,2), (1,4,3), (1,3), (1,2,4), (1,2,3,4), (1,4), (1,3,4)\},$$

and $|A(T)\phi A(S)| = 16$. Since we found a double coset of cardinality 16, we only have two double cosets, i.e two non-isomorphic tanglegrams corresponding to the trees $T$ and $S$.

# Chapter 2

# A formula for the enumeration of tanglegrams

Here, we will use the properties of trees and tanglegrams given in the previous chapter to derive a formula for the number of non-isomorphic tanglegrams with $n$ leaves. We note that this result was established by Billey, Konvalinka and Matsen in their paper [23]. Moreover, we can generalize the concept of tanglegrams by taking multiple binary trees and obtain new combinatorial objects called *tangled chains*. Then, the formula for tanglegrams can easily be extended to tangled chains. We first talk about binary partitions and their relation to the types of permutations in a binary tree $T$. Afterwards, we state and prove the formula for the number of non-isomorphic tanglegrams with $n$ leaves. Finally, we generalize this formula to tangled chains and give an asymptotic expansion for the number of tanglegrams of size $n$.

## 2.1 Binary partitions

Counting non-isomorphic tanglegrams involves counting non-isomorphic binary trees. Hence, we need to consider automorphism groups of binary trees with $n$ leaves which are subgroups of $S_n$. More precisely, we have to look at the types of permutations present in the automorphism group of a binary tree $T$. We will see that the type of a permutation in a binary tree $T$ is a binary partition.

First, let us recall some useful facts about permutations. It is well known that every permutation $\sigma \in S_n$ can be written as a product of disjoint cycles. Then, we define the type of a permutation $\sigma \in S_n$ to be the sequence of positive integers $\lambda = (k^{i_k})$, written in decreasing order with respect to $k$, where $i_k$ is the number of cycles of length $k$ in the disjoint decomposition of $\sigma$ into cycles.

For instance, the permutation $\sigma = (1,2)(3,4)(5,6,7) \in S_7$ is of type $(3^1, 2^2)$.

Let $G$ be a group. For two elements $a, b \in G$, we write $a \sim b$ if and only if there exists $c \in G$ such that $b = cac^{-1}$. The relation $\sim$ is an equivalence relation, and $b$ is called a *conjugate* of $a$. The next lemma is useful for characterising conjugates of an element in $S_n$.

**Lemma 2.1.1.** If $\sigma, \tau \in S_n$ such that $\sigma$ is a cycle i.e. $\sigma = (a_1, a_2, \ldots, a_s)$ $(s \leq n)$, then $\tau\sigma\tau^{-1} = (\tau(a_1), \tau(a_2), \ldots, \tau(a_s))$.

*Proof.* If $x \notin \{\tau(a_1), \tau(a_2), \ldots, \tau(a_s)\}$ then $\tau^{-1}(x) \notin \{a_1, \ldots, a_s\}$, so $\tau\sigma\tau^{-1}(x) = \tau\tau^{-1}(x) = x$. Otherwise, if $x = \tau(a_i)$, then $\tau\sigma\tau^{-1}(x) = \tau(\sigma(a_i)) = \tau(a_{i+1})$ and $\tau\sigma\tau^{-1}(\tau(a_s)) = \tau(a_1)$. $\square$

In the case of permutation groups, we have the following interesting proposition relating conjugacy and type of permutations.

**Proposition 2.1.2.** Two permutations $\sigma$ and $\sigma'$ are conjugate in $S_n$ if and only if they have the same type.

*Proof.* If $\sigma' = \tau\sigma\tau^{-1}$ and $\sigma = c_1 c_2 \ldots c_s$ is the decomposition of $\sigma$ into disjoint cycles then $\sigma' = \tau c_1 \tau^{-1} \tau c_2 \tau^{-1} \ldots \tau c_s \tau^{-1}$ and by Lemma 2.1.1, $c_i$ and $\tau c_i \tau^{-1}$ are of the same type. Conversely, if $\sigma = (a_{11}, \ldots, a_{1i_1}) \ldots (a_{r1}, \ldots, a_{ri_r})$ and $\sigma' = (a'_{11}, \ldots, a'_{1i_1}) \ldots (a'_{r1}, \ldots, a'_{ri_r})$, then we just have to take $\tau(a_{ik}) = a'_{ik}$ to obtain $\sigma' = \tau\sigma\tau^{-1}$. $\square$

Permutations of the same type play important roles in the enumeration of tanglegrams. More precisely, the number of permutations of a given type will appear in the formula for the number of tanglegrams of size $n$. This number involves the type of the given permutation itself as the next proposition asserts.

**Proposition 2.1.3.** Given a permutation $\sigma \in S_n$ of type $\lambda = (k^{i_k})$, the number of permutations which have the same type as $\sigma$ is given by $n!/z_\lambda$, where

$$z_\lambda = \prod_{1 \leq k \leq n} k^{i_k} i_k!.$$

*Proof.* By Proposition 2.1.2, the permutations that have the same type are conjugate in $S_n$. Let $\sigma \in S_n$ with type $\lambda = (k^{i_k})$. We will enumerate the conjugates of $\sigma$ using a constructive method.
Suppose $i_k \neq 0$, then in the disjoint decomposition of $\sigma$ into cycles, a product of $k$-cycles appears:

$$(a_{11}, \cdots, a_{1k})(a_{21}, \cdots, a_{2k}) \cdots (a_{i_k 1}, \cdots, a_{i_k k}).$$

If $c_i = (a_{i1}, \cdots, a_{ik})$ and $\tau \in S_n$, then

$$\begin{aligned} \tau c_1 c_2 \cdots c_{i_k} \tau^{-1} &= \tau c_1 \tau^{-1} \tau c_2 \cdots \tau c_{i_k} \tau^{-1} \\ &= (\tau(a_{11}), \cdots, \tau(a_{1k})) \cdots (\tau(a_{i_k 1}), \cdots, \tau(a_{i_k k})). \end{aligned}$$

Thus, the map $\sigma \mapsto \tau\sigma\tau^{-1}$ sends a cycle of length $k$ to a cycle of length $k$, and since all the cycles are disjoint , we have $i_k!$ ways of choosing the images of all cycles of length $k$. Now, suppose that $\tau(c_i) = c_l$. Then we have $k$ choices for $\tau(a_{i1})$ in $\{a_{l1}, \ldots, a_{lk}\}$. Once $\tau(a_{i1})$ is chosen, we do not have a choice for the other $\tau(a_{ij})$. Indeed, if $\tau(a_{i1}) = \tau(a_{lj})$

then $\tau(a_{i2}) = \tau(a_{lm})$ with $m \equiv j+1 \mod (k)$ and so on. Hence, we have $k^{i_k}$ choices for the image of $a_{i1}$ and $k^{i_k} \cdot i_k!$ choices for the image of $c_{i_k}$ by $\tau$. Since the $k^{i_k} \cdot i_k!$ choices are independent for each value $k$ with $i_k \neq 0$, the product $z_\lambda = \prod_{1 \leq k \leq n} k^{i_k} i_k!$ is the number of possible constructions of permutations of type $\lambda$ for any given sequence $(a_{ij})$. Now, the values $a_{ij}$ can range in $\{1, \ldots, n\}$ and they are all distinct so we have $n!$ ways of assigning values to the $a_{ij}$. Thus, the number of permutations that have the same type as $\sigma$ is $n!/z_\lambda$. $\square$

Next, we will relate types of permutations to partitions. Here, partitions are defined as follows.

**Definition 2.1.4.** A partition is a weakly decreasing sequence of positive integers $(\lambda_1, \lambda_2, \ldots, \lambda_k)$. We say that a partition is a *binary partition* if all the parts of the partition are integer powers of two.

Consider an element $\sigma$ of $S_n$ of type $(k^{i_k})$. The sequence $(k^{i_k})$ is identified to a partition of $n$ where $k$ is repeated $i_k$ times. For instance, $\sigma = (3,4,5) \in S_5$ is of type $(3, 1^2)$, and the corresponding partition is given by $(3,1,1)$. From here we refer to the type of a permutation as a partition. In some cases, for convenience, we will omit parenthesis and commas when we write a partition.

We will also adopt the following operations on partitions. The union of two partitions is the partition obtained by combining all the parts of the two partitions. Multiplying or dividing a partition by an integer $\alpha$ is equivalent to multiplying or dividing each part by $\alpha$. For example, $(4,2,2,1) \cup (5,3,3) = (5,4,3,3,2,2,1)$ and $2 \cdot (3,1,1) = (6,2,2)$.

The generating functions for binary partitions of different types have been established by Sloane and Seller in their paper [30]. Those generating functions might play a role in establishing the generating function for non-isomorphic tanglegrams, which is still an open problem.

We recursively define a linear order on the set of binary trees. This linear order will be useful for the proof of the exact formula for the number of tanglegrams of size $n$.

**Definition 2.1.5.** Let $T$ and $S$ be two binary trees. We say that $T > S$ if

- $T$ has more leaves than $S$ or
- $T$ and $S$ have the save number of leaves, $T$ has subtrees $T_1$ and $T_2$, $T_1 \geq T_2$, $S$ has subtrees $S_1$ and $S_2$, $S_1 \geq S_2$, and
  - $T_1 > S_1$ or
  - $T_1 = S_1$ and $T_2 > S_2$.

We say that $T$ and $S$ are equal $(T = S)$ if neither $T < S$ nor $T > S$.

Then, we have the following proposition:

**Proposition 2.1.6** ([3])**.** Two binary trees $T$ and $S$ are equal if and only if $T$ is isomorphic to $S$.

*Proof.* ($\Rightarrow$) We suppose that $T$ and $S$ are equal. The proof will be by induction on the number of leaves $n$ of $T$ and $S$. For $n = 1$, the two binary trees $T$ and $S$ are both the tree with one leaf so they are isomorphic. Suppose that the statement is true for all pairs of binary trees $(T, S)$ which have a number of leaves less than or equal to $n$ and satisfying $T = S$. Now, suppose $T$,$S$ have $n + 1$ leaves. Let $T_1, T_2$ be the branches of $T$ and $S_1, S_2$ be the branches of $S$. Since $T = S$, $T_1$ is equal to one of $S_1$ and $S_2$, the same goes for $T_2$. We can assume without loss of generality that $T_1 = S_1$ and that $T_2 = S_2$. By induction, $T_1$ is isomorphic to $S_1$ and $T_2$ is isomorphic to $S_2$. Let $\phi_1 : L(T_1) \to L(S_1)$, $\phi_2 : L(T_2) \to L(S_2)$ be isomorphisms of $T_1$ and $S_1$ ($T_2$ and $S_2$ respectively). Then, the map

$$\phi(x) = \begin{cases} \phi_1(x) & \text{if } x \in L(T_1) \\ \phi_2(x) & \text{if } x \in L(T_2) \end{cases}$$

is an isomorphism from $T$ to $S$.

($\Leftarrow$) We suppose that $T$ is isomorphic to $S$. We proceed again by induction on the number of leaves $n$ of $T$ and $S$. For $n = 1$, the statement is true since $T$ and $S$ are the tree with one leaf. Suppose that the statement is true for all pairs of isomorphic trees $T$ and $S$ with a number of leaves less than or equal to $n$. Now, suppose $T$,$S$ have $n + 1$ leaves. Let $T_1, T_2$ be the branches of $T$ and $S_1, S_2$ be the branches of $S$. Since $T$ is isomorphic to $S$, $T_1$ is isomorphic to $S_1$ or $S_2$, the same goes for $T_2$. Assume that $T_1$ is isomorphic to $S_1$ and $T_2$ is isomorphic to $S_2$. By induction $T_1 = S_1$ and $T_2 = S_2$. It follows that $T = S$.

□

Now, let $B_n$ be the set of all non-plane binary trees with $n$ leaves and $T \in B_n$. We label the leaves of $T$ by the numbers $1, \dots, n$ in order to define the automorphism group. We recall from Proposition 1.1.2 that an automorphism $\sigma$ of $T$ is identified by the bijection on the set of leaves. Let $T_1$ and $T_2$ be the two branches of $T$. From Proposition 1.1.6, we know that if $T$ has two branches $T_1$ and $T_2$ then the automorphism group $A(T)$ of $T$ is isomorphic to $A(T_1) \times A(T_2)$ if $T_1 \neq T_2$ and to $A(T_1) \wr \mathbb{Z}_2$ if $T_1 = T_2$. In addition, $A(T)$ can be obtained from copies of $\mathbb{Z}_2$ by direct and wreath products. The proposition below links binary partitions to types of elements in $A(T)$.

**Proposition 2.1.7** ([3]). Let $T \in B_n$ and let $T_1$ and $T_2$ be the branches of $T$. We have the following properties:

(1) If $T_1 \neq T_2$ then a permutation $\sigma$ in $A(T)$ is of type $\lambda = \lambda^1 \cup \lambda^2$ where $\lambda^i$ is the type of an element of $A(T_i)$, $i = 1, 2$.

(2) If $T_1 = T_2$ then we have two cases for the type of a permutation $\sigma$ in $A(T)$:

  (a) $\lambda = \lambda^1 \cup \lambda^2$ or

  (b) $\lambda = 2\lambda^1$ where $\lambda^i$ is the type of an element of $A(T_i)$, $i = 1, 2$.

(3) The type of an element of $A(T)$ is a binary partition.

*Proof.* We label the leaves of $T$ by the numbers $1, \ldots, n$ in such a way that the labels of $T_1$ are from 1 to $k$ and the labels of $T_2$ are from $k+1$ to $n$. Consider each $A(T_i)$ to be a subgroup of the permutations of the leaf labels for $T_i$. More precisely, the automorphism group of $T_1$ will be on the set of labels $\{1, \ldots, k\}$ and the automorphism group of $T_2$ will be on the set of labels $\{k+1, \ldots, n\}$. A pair $(\sigma_1, \sigma_2)$ of elements in $A(T_1) \times A(T_2)$ corresponds to an element of $A(T)$ which fixes the elements of the set $\{k+1, \ldots, n\}$ ($\{1, \ldots, k\}$ respectively).

For Part (1) of the proposition, if $T_1 \neq T_2$ then an element $\sigma$ of $A(T)$ is the product of an element $\sigma_1 \in A(T_1)$ and an element $\sigma_2 \in A(T_2)$. So, if $\sigma_1$ is of type $\lambda^1$ and $\sigma_2$ is of type $\lambda^2$ then $\lambda$ is of type $\lambda^1 \cup \lambda^2$ (since all the cycles in $\sigma_1$ are disjoint from all the cycles in $\sigma_2$).

For Part (2) of the proposition, assume that $T_1 = T_2$ and let $\sigma \in A(T)$. If $\sigma$ sends $L(T_1)$ to $L(T_1)$, then $\sigma$ must send $L(T_2)$ to $L(T_2)$. So, $\sigma$ can be written as a disjoint product $\sigma_1\sigma_2$ where $\sigma_1 \in A(T_1)$ and $\sigma_2 \in A(T_2)$. Thus, $\sigma$ has type $\lambda = \lambda^1 \cup \lambda^2$ where $\lambda^1$ is the type of $\sigma_1$ and $\lambda^2$ is the type of $\sigma_2$. This gives us Part (2) (a) of Proposition 2.1.7.

Now, if $\sigma$ sends $L(T_1)$ to $L(T_2)$ (so $L(T_2)$ is sent to $L(T_1)$) then $\sigma^2$ must send $L(T_1)$ to $L(T_1)$ and $L(T_2)$ to $L(T_2)$. Therefore, $\sigma^2 = \sigma_1\sigma_2$ where $\sigma_1 \in A(T_1)$ and $\sigma_2 \in A(T_2)$. Since $\sigma$ sends leaves from $T_1$ to $T_2$ and vice versa, all cycles of $\sigma$ must be of even length (in the disjoint cycle decomposition). Indeed, if there is a cycle of odd length, say $(a_1\ a_2\ \ldots\ a_{2l+1})$ in $\sigma$ then $a_1$ and $a_{2l+1}$ must be leaves of the same branch, which is a contradiction since $\sigma(a_{2l+1}) = a_1$. Next, let $(a_1\ a_2 \ldots\ a_{2l})$ be a cycle in $\sigma$, we can assume without loss of generality that $a_1 \in L(T_1)$. Then,

$$(a_1\ a_2\ \cdots\ a_{2l})^2 = (a_1\ a_3\ \cdots\ a_{l+1})(a_2\ a_4\ \cdots\ a_{2l})$$

where $(a_1\ a_3\ \cdots\ a_{l+1})$ and $(a_2\ a_4\ \cdots\ a_{2l})$ are cycles of $\sigma_1$ and $\sigma_2$ respectively of the same length $l$. This implies that

- first, if there are $i_l$ cycles of length $l$ in $\sigma_1$ then there are $i_l$ cycles of length $l$ in $\sigma_2$ which means that the permutations $\sigma_1$ and $\sigma_2$ have the same type, say $\lambda^1$,
- second, a cycle of length $2l$ in $\sigma$ splits to two cycles of length $l$ in $\sigma_1$ and $\sigma_2$.

This last two fact imply that $\sigma$ is of type $2\lambda^1$. This gives us Part (2) (b) of Proposition 2.1.7
Finally, since the automorphism group of the tree with one vertex is trivial, property (1) and (2) will imply property (3) by induction.

□

For two binary trees $T$ and $S$, $A(T)_\lambda$ and $A(S)_\lambda$ are the sets of permutations of $A(T)$ and $A(S)$ of type $\lambda$ respectively. The next proposition gives information about the size of $A(T)_\lambda$ for a given $\lambda$.

**Proposition 2.1.8.** For a binary tree $T$ with root branches $T_1$ and $T_2$, and a binary partition $\lambda$ we have the following cases:

- if $T_1 \neq T_2$, then
$$|A(T)_\lambda| = \sum_{\lambda=\lambda^1\cup\lambda^2} |A(T_1)_{\lambda^1}|\,|A(T_2)_{\lambda^2}|\,,$$
- if $T_1 = T_2$, then
$$|A(T)_\lambda| = \sum_{\lambda=\lambda^1\cup\lambda^2} |A(T_1)_{\lambda^1}|\,|A(T_1)_{\lambda^2}| + |A(T_1)_{\lambda/2}|\,|A(T_1)|\,.$$

*Proof.* By the arguments in the proof of the previous proposition, the first case is clear and so is the sum in the second case. It remains to prove the additional term in the second case. This term should give the number of automorphisms of type $\lambda$ of $T$ that swap $T_1$ and $T_2$. Assume that $T_1 = T_2$, so $n$ is even, say $n = 2k$. We label the leaves of $T_1$ with $\{1, 2, \ldots, k\}$, and the leaves of $T_2$ with $\{k+1, k+2, \ldots, k+k\}$ in such a way that

$$\pi = (1, k+1)(2, k+2)\cdots(k, k+k) \in A(T).$$

For a $\sigma_1 \in A(T_1)_{\lambda/2}$ and a $\sigma_2 \in A(T_2)$, we can construct an element $\sigma$ of $A(T)_\lambda$ in the following way:

$$\sigma = \sigma_2\sigma_1\pi\sigma_2^{-1}.$$

It is straightforward to check that $\sigma$ sends $T_1$ to $T_2$ and vice versa. Moreover, $\sigma^2$ acts on $T_1$ like $\sigma_1$ does (i.e. $\sigma^2(j) = \sigma_1(j)$ for $j \in \{1, 2, \ldots, k\}$). Therefore, again as in the proof of the previous proposition, the type of $\sigma$ must be $\lambda$ (twice the type of $\sigma_1$).

The above construction can be reversed i.e. given $\sigma \in A(T)_\lambda$ that sends $T_1$ to $T_2$, we can recover $\sigma_1$ and $\sigma_2$ by the following formulas:

$$\sigma_1(j) = \sigma^2(j) \quad \text{and} \quad \sigma_2(k+j) = \sigma(j),$$

for $j \in \{1, 2, \ldots, k\}$. Hence, the number of automorphisms of type $\lambda$ of $T$ that swap $T_1$ and $T_2$ is

$$|A(T_1)_{\lambda/2}|\,|A(T_2)| = |A(T_1)_{\lambda/2}|\,|A(T_1)|\,.$$

This completes the proof. □

## 2.2 Exact enumeration of non-isomorphic tanglegrams

We are now ready to prove the main theorem of this chapter, which states as the following.

**Theorem 2.2.1** ([3])**.** *The number $t_n$ of non-isomorphic tanglegrams with $n$ leaves is given by*

$$t_n = \sum_{\lambda} \frac{\prod_{i=2}^{l(\lambda)}(2(\lambda_i + \cdots + \lambda_{l(\lambda)}) - 1)^2}{z_\lambda}, \tag{2.2.1}$$

*where the sum is taken over all binary partition $\lambda = (\lambda_1, \lambda_2, \ldots)$ of $n$, $l(\lambda)$ is the length of $\lambda$ i.e. the number of parts, and $z_\lambda$ is defined as in Proposition 2.1.3.*

For instance, the binary partitions of $n = 3$ are $(2,1)$ and $(1,1,1)$, so the number of non-isomorphic tanglegrams of size 3 is given by

$$t_3 = \frac{1^2}{2} + \frac{3^2}{6} = 2.$$

The first 10 terms of the sequence $t_n$ starting at $n = 1$ are

$$1, 1, 2, 13, 114, 1509, 25595, 535753, 13305590, 382728552,$$

see [31, A258620] for more terms.
In order to prove Theorem 2.2.1, we first need some auxiliary results.

**Proposition 2.2.2** ([3])**.** For a binary partition $\lambda$,

$$\sum_{T \in B_n} \frac{|A(T)_\lambda|}{|A(T)|} = \frac{\prod_{i=2}^{l(\lambda)} (2(\lambda_i + \cdots + \lambda_{l(\lambda)}) - 1)}{z_\lambda}, \tag{2.2.2}$$

where $A(T)_\lambda$ denotes the elements of $A(T)$ of type $\lambda$.

We remark here that the formula given in Proposition 2.2.2 implies the following theorem:

**Theorem 2.2.3** ([3])**.** *The number $b_n$ of non-isomorphic binary trees with $n$ leaves is given by:*

$$b_n = \sum_{\lambda} \frac{\prod_{i=2}^{l(\lambda)} (2(\lambda_i + \cdots + \lambda_{l(\lambda)}) - 1)}{z_\lambda}, \tag{2.2.3}$$

*where the sum is taken over all binary partition $\lambda = (\lambda_1, \lambda_2, \dots)$ of $n$.*

*Proof.* We have

$$b_n = \sum_{T \in B_n} 1 = \sum_{T \in B_n} \sum_{\lambda} \frac{|A(T)_\lambda|}{|A(T)|} = \sum_{\lambda} \sum_{T \in B_n} \frac{|A(T)_\lambda|}{|A(T)|}.$$

Thus, by Proposition 2.2.2,

$$b_n = \sum_{\lambda} \frac{\prod_{i=2}^{l(\lambda)} (2(\lambda_i + \cdots + \lambda_{l(\lambda)}) - 1)}{z_\lambda}.$$

□

Theorem 2.2.3 gives a new formula for the number of non-isomorphic binary trees. These trees are enumerated by the Wedderburn-Etherington numbers, whose sequence starts with

$$0, 1, 1, 1, 2, 3, 6, 11, 23, 46, 98, 207, 451, 983, 2179, 4850,$$

see ([31, A001190]) for more terms.
In order to prove Proposition 2.2.2, we will need a recurrence relation involving the partition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_{l(\lambda)})$. This is established in the following lemma. For a nonempty subset $S = \{i_1 < i_2 < \cdots < i_k\}$ of the natural numbers, define

$$r_S(x_1, x_2, \cdots) = (x_{i_2} + x_{i_3} + \cdots + x_{i_k} - 1) \cdots (x_{i_{k-1}} + x_{i_k} - 1)(x_{i_k} - 1). \tag{2.2.4}$$

Let $x$ denotes the sequence $(x_1, x_2, \dots)$ and $x/2$ denotes the sequence $(x_1/2, x_2/2, \dots)$. The lemma states as follows:

**Lemma 2.2.4** ([3])**.** Let $n \geq 2$, then

$$r_{[\![n]\!]}(x) = 2^{n-1} r_{[\![n]\!]}(x/2) + \sum_{1 \in S \subsetneq [\![n]\!]} r_S(x) \cdot r_{[\![n]\!] \setminus S}(x), \tag{2.2.5}$$

where $[\![n]\!]$ denotes the set $\{1, \dots, n\}$.

For example, for $n = 3$, we have

$$r_{[\![3]\!]}(x) = (x_2 + x_3 - 1)(x_3 - 1) = (x_2 + x_3 - 2)(x_3 - 2) + 1 \cdot (x_3 - 1) + (x_2 - 1) \cdot 1 + (x_3 - 1) \cdot 1,$$

where the last three terms on the right hand side correspond respectively to the subsets $\{1\}$, $\{1, 2\}$, $\{1, 3\}$.
From here, for a polynomial or a power series $p(t)$, $[t^n]p(t)$ denotes the coefficient of $t^n$ in $p(t)$. If $t = (t_1, t_2, \dots, t_k)$ then $[t_i^n]p(t)$ denotes the coefficient of $t_i^n$ in $p(t)$.

*Proof of Lemma 2.2.4.* The proof is by induction on $n$. For $n = 2$, we have $r_{[\![2]\!]}(x) = (x_2 - 2) + 1 \cdot 1$ so the statement is true. Assume that the statement is true for every natural number $k$ such that $k \leq n - 1$, we will prove it for $n$. We have

$$\begin{aligned} r_{[\![n]\!]}(x) &= (x_2 + x_3 + \dots + x_n - 1)(x_3 + \dots + x_n - 1) \cdots (x_n - 1) \\ &= (x_2 + x_3 + \dots + x_n - 1) r_{[\![2,n]\!]}(x). \end{aligned}$$

So $r_{[\![n]\!]}(x)$ is linear with respect to $x_2$ and the coefficient of $x_2$ in $r_{[\![n]\!]}(x)$ is $r_{[\![2,n]\!]}(x)$. We remark here that for any real non negative numbers $a, b$, $[\![a, b]\!]$ denotes the set of natural numbers (possibly empty) in the interval $[a, b]$. Furthermore,

$$\begin{aligned} 2^{n-1} r_{[\![n]\!]}(x/2) &= (x_2 + x_3 + \dots + x_n - 2)(x_3 + \dots + x_n - 2) \cdots (x_n - 2) \\ &= (x_2 + x_3 + \dots + x_n - 2) 2^{n-2} r_{[\![2,n]\!]}(x/2). \end{aligned}$$

So $2^{n-1} r_{[\![n]\!]}(x/2)$ is also linear with respect to $x_2$ and the coefficient of $x_2$ in $2^{n-1} r_{[\![n]\!]}(x/2)$ is $2^{n-2} r_{[\![2,n]\!]}(x/2)$.
Next, we note that $r_S(x) \cdot r_{[\![n]\!] \setminus S}(x)$ contains $x_2$ if and only if $2 \in S$. Indeed, since $S$ contains 1 and $1 < 2$, if $2 \in S$ then $x_2$ appears in $r_S(x)$. Thus $r_S(x) \cdot r_{[\![n]\!] \setminus S}(x)$ contains $x_2$. Conversely, suppose that $r_S(x) \cdot r_{[\![n]\!] \setminus S}(x)$ contains $x_2$. Assume first that $2 \in [\![2, n]\!] \setminus S$ (thus $2 \notin S$). Then 2 is the minimum value in $[\![2, n]\!]$, so $r_{[\![2,n]\!] \setminus S}$ does not contain $x_2$. Since $2 \notin S$, $x_2$ does not appear in $r_S(x)$ contradicting the assumption that $r_S(x) \cdot r_{[\![n]\!] \setminus S}(x)$ contains $x_2$. Hence $2 \in S$. If $S = \{1, 2, j_1 < j_2 \cdots < j_k\}$, where $j_i \neq 1, 2$ for $i \in \{1, 2, \cdots, k\}$, then

$$\begin{aligned} r_S(x) &= (x_2 + x_{j_1} + \dots + x_{j_k} - 1)(x_{j_1} + \dots + x_{j_k} - 1) \cdots (x_{j_k} - 1) \\ &= (x_2 + x_{j_1} + \dots + x_{j_k} - 1) r_{S \setminus \{1\}}(x) \\ &= x_2 \cdot r_{S \setminus \{1\}}(x) + (x_{j_1} + \dots + x_{j_k} - 1) \cdot r_{S \setminus \{1\}}(x). \end{aligned}$$

Consequently, $r_S(x)$ is linear in $x_2$ and so is $r_S(x) \cdot r_{[\![n]\!]\setminus S}(x)$. In addition, we notice that if $1, 2 \in S$, then $[\![n]\!] \setminus S = [\![2,n]\!] \setminus S$. So

$$[x_2] r_S(x) \cdot r_{[\![n]\!]\setminus S}(x) = r_{S\setminus\{1\}}(x) \cdot r_{[\![n]\!]\setminus S}(x) = r_{S\setminus\{1\}}(x) \cdot r_{[\![2,n]\!]\setminus S}(x) = r_{S'}(x) \cdot r_{[\![2,n]\!]\setminus S'}(x),$$

where $S' = S \setminus \{1\}$. Hence both sides of Equation (2.2.5) are linear with respect to $x_2$. Thus to prove they are equal, it is sufficient to prove that they have the same coefficient for $x_2$ and that they are the same for one value of $x_2$. By the induction hypothesis,

$$r_{[\![n-1]\!]}(x) = 2^{n-2} r_{[\![n-1]\!]}(x/2) + \sum_{1\in S\subsetneq[\![n-1]\!]} r_S(x) \cdot r_{[\![n-1]\!]\setminus S}(x).$$

Since the function $g : [\![n-1]\!] \to [\![2,n]\!]$ defined by $g(i) = i+1$ is a bijection, from the previous relation we have

$$r_{[\![2,n]\!]}(x) = 2^{n-2} r_{[\![2,n]\!]}(x/2) + \sum_{2\in S\subsetneq[\![2,n]\!]} r_S(x) \cdot r_{[\![2,n]\!]\setminus S}(x),$$

so the left and right hand side of Equation (2.2.5) have the same $x_2$ coefficients.
Now, we plug the value $x_2 = 2 - x_3 - x_4 - \cdots - x_n$ into $r_{[\![n]\!]}(x)$. The first factor $(x_2 + x_3 + \cdots + x_n - 1)$ of the product in $r_{[\![n]\!]}(x)$ disappears and the left hand side of (2.2.5) becomes $r_{[\![n]\!]\setminus\{2\}}(x)$. On the right hand side, $r_{[\![n]\!]}(x/2) = 0$ since the first factor becomes zero after plugging in $2 - x_3 - x_4 - \cdots - x_n$ for $x_2$. Assume that $S = \{1, 2, x_{i_1}, \cdots, x_{i_k}\}$ and $[\![n]\!] \setminus S = \{x_{j_1}, \cdots, x_{j_p}\}$ where $j_l \neq 1, 2$ for $l$. After we plug in the value $2 - x_3 - x_4 - \cdots - x_n$ for $x_2$, we have

$$r_S(x) = -(x_{j_1} + x_{j_2} + \cdots + x_{j_p} - 1)(x_{i_1} + x_{i_2} + \cdots + x_{i_k} - 1)\cdots(x_{i_k} - 1),$$

and

$$r_{[\![n]\!]\setminus S}(x) = (x_{j_2} + x_{j_3} + \cdots + x_{j_p} - 1)\cdots(x_{i_{k-1}} + x_{i_k} - 1)(x_{i_k} - 1).$$

Furthermore, we have

$$r_{S\setminus\{2\}}(x) = (x_{i_1} + x_{i_2} + \cdots + x_{i_k} - 1)\cdots(x_{i_k} - 1)$$

and

$$r_{([\![n]\!]\setminus S)\cup\{2\}}(x) = (x_{j_1} + x_{j_2} + \cdots + x_{j_p} - 1)\cdots(x_{i_{k-1}} + x_{i_k} - 1)(x_{i_k} - 1).$$

Thus, $r_S(x) \cdot r_{[\![n]\!]\setminus S}(x) + r_{S\setminus\{2\}}(x) \cdot r_{([\![n]\!]\setminus S)\cup\{2\}}(x) = 0$. All the term in the summation cancel except $r_{[\![n]\!]\setminus\{2\}}(x) \cdot r_{\{2\}}(x) = r_{[\![n]\!]\setminus\{2\}}(x)$, thus the right hand side of Equation (2.2.5) is equal to the left hand side. $\square$

Once Lemma 2.2.4 is proven, we can proceed to the proof of Proposition 2.2.2. Recall that for two binary trees $T$ and $S$, $A(T)_\lambda$ and $A(S)_\lambda$ are the permutations of $A(T)$ and $A(S)$ of type $\lambda$ respectively.

*Proof of Proposition 2.2.2.* Suppose $\lambda$ is a binary partition of $n$. The proof is by induction on $n$. For $n = 1$, we have only one tree $T$ which is the one leaf tree and one partition of $n$ which is $\lambda = (1)$. Hence,

$$\sum_{T \in B_1} \frac{|A(T)_\lambda|}{|A(T)|} = 1.$$

Also, since $\lambda = (1)$, $\prod_{i=2}^{l(\lambda)}(2(\lambda_i + \cdots + \lambda_{l(\lambda)}) - 1) = 1$ and $z_\lambda = 1$. Hence, the statement is true for $n = 1$.

Now, assume that Equation (2.2.2) is true for all $k \leq n - 1$. We look for the case $k = n$. First, we need to differentiate between the case where the branches $T_1$ and $T_2$ of $T$ are different and when they are equal. We can assume without loss of generality that $T_1 > T_2$ if $T_1 \neq T_2$, so

$$\sum_{T \in B_n} \frac{|A(T)_\lambda|}{|A(T)|} = \sum_{T_1 > T_2} \frac{|A(T)_\lambda|}{|A(T)|} + \sum_{T_1 = T_2} \frac{|A(T)_\lambda|}{|A(T)|}. \tag{2.2.6}$$

Recall that if $T_1 = T_2$ then $|A(T)| = |A(T_1) \wr \mathbb{Z}_2| = 2|A(T_1)|^2$ and $|A(T)| = |A(T_1) \times A(T_2)| = |A(T_1)| \cdot |A(T_2)|$ if $T_1 \neq T_2$. Moreover, from Proposition 2.1.7 we know that if $T_1 > T_2$ then a permutation $\sigma$ in $A(T)$ is of type $\lambda = \lambda^1 \cup \lambda^2$ where $\lambda^i$ is the type of an element of $A(T_i)$, $i = 1, 2$. However, if $T_1 = T_2$ then there are two possible types for a permutation $\sigma \in A(T)$: $\lambda = \lambda^1 \cup \lambda^2$ and $\lambda = 2\lambda^1$. So, by the previous observations and Proposition 2.1.8, Equation (2.2.6) splits in the following way:

$$\begin{aligned}\sum_{T \in B_n} \frac{|A(T)_\lambda|}{|A(T)|} = &\sum_{T_1 > T_2} \Big( \sum_{\lambda = \lambda^1 \cup \lambda^2} \frac{|A(T_1)_{\lambda^1}| \cdot |A(T_2)_{\lambda^2}|}{|A(T_1)| \cdot |A(T_2)|} \Big) \\ &+ \sum_{T_1} \frac{\sum_{\lambda = \lambda^1 \cup \lambda^2} |A(T_1)_{\lambda^1}| \cdot |A(T_1)_{\lambda^2}| + |A(T_1)| \cdot |A(T_1)_{\lambda/2}|}{2|A(T_1)|^2}.\end{aligned}$$

Equivalently,

$$\begin{aligned}2\sum_{T \in B_n} \frac{|A(T)_\lambda|}{|A(T)|} = &\sum_{T_1 \in B_{n/2}} \frac{|A(T_1)_{\lambda/2}|}{|A(T_1)|} \\ &+ \sum_{\lambda = \lambda^1 \cup \lambda^2} \Big( \sum_{T_1 \in B_{|\lambda^1|}} \frac{|A(T_1)_{\lambda^1}|}{|A(T_1)|} \Big) \Big( \sum_{T_2 \in B_{|\lambda^2|}} \frac{|A(T_2)_{\lambda^2}|}{|A(T_2)|} \Big)\end{aligned} \tag{2.2.7}$$

where $|\lambda^i|$ is the sum of all parts of $\lambda^i$ for $i = 1, 2$. Now, we define

$$R_\lambda = \frac{\prod_{i=2}^{l(\lambda)}(2(\lambda_i + \lambda_2 + \cdots + \lambda_{l(\lambda)}) - 1)}{z_\lambda} = \frac{r_{[\![l(\lambda)]\!]}(2\lambda_1, 2\lambda_2, 2\lambda_3, \cdots)}{z_\lambda}.$$

By the induction hypothesis, the right hand side of Equation (2.2.7) is

$$R_{\lambda/2} + \sum_{\lambda = \lambda^1 \cup \lambda^2} R_{\lambda^1} \cdot R_{\lambda^2}.$$

It remains to check that

$$2R_\lambda = R_{\lambda/2} + \sum_{\lambda=\lambda^1\cup\lambda^2} R_{\lambda^1}\cdot R_{\lambda^2}. \tag{2.2.8}$$

We note that if $\lambda = 2\lambda^1$ then $z_\lambda = 2^{l(\lambda)} z_{\lambda/2}$. So, mutltiplying both sides of Equation (2.2.8) by $z_\lambda$ gives

$$\begin{aligned}2\prod_{i=2}^{l(\lambda)}(2(\lambda_i+\cdots+\lambda_{l(\lambda)})-1) &= 2^{l(\lambda)}\prod_{i=2}^{l(\lambda)}(\lambda_i+\cdots+\lambda_{l(\lambda)}-1)\\ &+\sum_{\lambda=\lambda^1\cup\lambda^2}\binom{\lambda}{\lambda^1,\lambda^2}\prod_{i=2}^{l(\lambda^1)}(2(\lambda_i^1+\cdots+\lambda^1_{l(\lambda^1)})-1)\cdot\prod_{i=2}^{l(\lambda^2)}(2(\lambda_i^2+\cdots+\lambda^2_{l(\lambda^2)})-1),\end{aligned}$$

where $\binom{\lambda}{\lambda^1,\lambda^2} = \prod_i \binom{m_i(\lambda)}{m_i(\lambda^1)}$ and $m_i(\lambda)$ ($m_i(\lambda^1)$ respectively) is the number of ocurrences of $2^i$ in the partition $\lambda$ ($\lambda^1$ respectively). Given a cycle type $\lambda$ and a cycle type $\lambda^1$, $\binom{\lambda}{\lambda^1,\lambda^2} = \prod_i \binom{m_i(\lambda)}{m_i(\lambda^1)}$ gives the number of ways of constructing a cycle type $\lambda^2$ such that $\lambda = \lambda^1 \cup \lambda^2$. The last equality holds by taking $x_i = 2\lambda_i$ in Lemma 2.2.4. This ends the proof of Proposition 2.2.2. □

Finally, we prove the main theorem of this chapter.

*Proof of Theorem 2.2.1.* From Proposition 1.2.4, we have

$$t_n = \sum_{T\in B_n}\sum_{S\in B_n} |\mathcal{C}(T,S)|,$$

where $\mathcal{C}(T,S)$ is the set of double cosets of $S_n$ with respect to $A(T)$ and $A(S)$.
Let us fix $T,S\in B_n$ and write $\mathcal{C} = \mathcal{C}(T,S)$, then

$$|\mathcal{C}| = \sum_{C\in\mathcal{C}} 1 = \sum_{C\in\mathcal{C}}\frac{|C|}{|C|} = \sum_{C\in\mathcal{C}}\sum_{\sigma\in C}\frac{1}{|C|}.$$

For all $\sigma\in S_n$, there exists a unique double coset (the equivalence class containing $\sigma$) $C_\sigma$ such that $\sigma\in C_\sigma$, so

$$|\mathcal{C}| = \sum_{\sigma\in S_n}\frac{1}{|C_\sigma|}.$$

From Proposition 1.2.8, we have

$$|C_\sigma| = \frac{|A(T)|\cdot|A(S)|}{|A(T)\cap\sigma A(S)\sigma^{-1}|}.$$

Consequently,

$$|\mathcal{C}| = \sum_{\sigma\in S_n}\frac{|A(T)\cap\sigma A(S)\sigma^{-1}|}{|A(T)|\cdot|A(S)|}.$$

We have

$$\sum_{\sigma \in S_n} |A(T) \cap \sigma A(S) \sigma^{-1}| = \sum_{\sigma \in S_n} \sum_{a \in A(T)} \sum_{b \in A(S)} I(a = \sigma b \sigma^{-1}),$$

where $I$ is the indicator function. Note that $a = \sigma b \sigma^{-1}$ can only be true if $a$ and $b$ are conjugate. By Proposition 2.1.2, $a$ and $b$ are conjugate if and only if they are of the same type $\lambda$. Moreover, the number of permutations $\sigma$ such that $a = \sigma b \sigma^{-1}$ is given by $z_\lambda$ (using the same idea as in the proof of Proposition 2.1.3). Thus,

$$\sum_{\sigma \in S_n} |A(T) \cap \sigma A(S) \sigma^{-1}| = \sum_{\lambda} |A(T)_\lambda| \cdot |A(S)_\lambda| \cdot z_\lambda$$

Thus,

$$|\mathcal{C}| = \frac{\sum_\lambda |A(T)_\lambda| \cdot |A(S)_\lambda| \cdot z_\lambda}{|A(T)| \cdot |A(S)|},$$

which implies that

$$\begin{aligned} t_n &= \sum_{T \in B_n} \sum_{S \in B_n} \sum_{\lambda} \frac{|A(T)_\lambda| \cdot |A(S)_\lambda| \cdot z_\lambda}{|A(T)| \cdot |A(S)|} \\ &= \sum_{\lambda} z_\lambda \sum_{T \in B_n} \sum_{S \in B_n} \frac{|A(T)_\lambda| \cdot |A(S)_\lambda|}{|A(T)||A(S)|} \\ &= \sum_{\lambda} z_\lambda \Big( \sum_{T \in B_n} \frac{|A(T)|_\lambda}{|A(T)|} \Big)^2. \end{aligned}$$

By Proposition 2.2.2,

$$\sum_{T \in B_n} \frac{|A(T)|_\lambda}{|A(T)|} = \frac{\prod_{i=2}^{l(\lambda)} (2(\lambda_i + \cdots + \lambda_{l(\lambda)}) - 1)}{z_\lambda},$$

so

$$t_n = \sum_{\lambda} \frac{\prod_{i=2}^{l(\lambda)} (2(\lambda_i + \cdots + \lambda_{l(\lambda)}) - 1)^2}{z_\lambda}.$$

□

Now, we look at a generalized version of tanglegrams called *tangled chains*.

**Definition 2.2.5** ([3]). Let $T_1, T_2, \ldots, T_p$ be binary trees. A *tangled chain* is a pair of tuples $((T_1, T_2, \ldots, T_p), (\phi_{ij})_{i,j \in \{1,\ldots,p\}})$ where $\phi_{ij} : L(T_i) \to L(T_j)$ such that the $\phi_{ij}$'s are bijections satisfying:

(1) $\phi_{ii} = Id$ for all $i$, (here $Id$ is the identity map from $L(T_i)$ to $L(T_i)$)

(2) $\phi_{ji} = \phi_{ij}^{-1}$ for all $i$, $j$ ,

(3) $\phi_{ik} = \phi_{ij} \circ \phi_{jk}$ for all $i$, $j$, $k$.

A tangled chain with 3 leaves and 3 binary trees is drawn in Figure 2.1, where the bijections are represented by inter-tree edges.
We can see that the $n^2$ bijections $\phi_{ij}$ are completely determined by the $n-1$ bijections $\{\phi_{1i}\}_{i=2,\ldots,n}$ since $\phi_{ij} = \phi_{1i}^{-1} \circ \phi_{1j}$ by property (2) and (3). Moreover, by property (3), we have

$$\phi_{1j} = \phi_{12} \circ \phi_{23} \cdots \circ \phi_{(j-1)j} ,$$

and

$$\phi_{1i}^{-1} = (\phi_{12} \circ \phi_{23} \cdots \circ \phi_{(i-1)i})^{-1},$$

for $i, j = 2, \ldots, n$. So the sequence $\phi_{12}, \phi_{23}, \ldots, \phi_{(p-1)p}$ also determines completely the bijections $\phi_{ij}$.

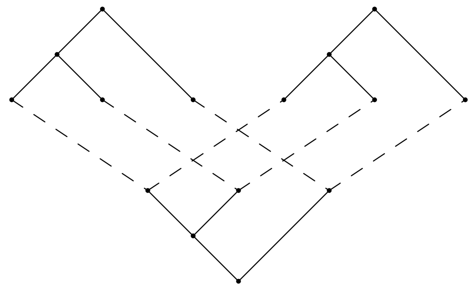

Figure 2.1: A tangled chain with 3 leaves and 3 binary trees.

As in the case of tanglegrams, the following definition tells us when two tangled chains are isomorphic.

**Definition 2.2.6** ([3]). Two tangled chains $X = ((T_1, T_2, \ldots, T_p), (\phi_{ij})_{i,j\in\{1,\ldots,p\}})$ and $X' = ((T_1, T_2, \cdots, T_p), (\phi'_{ij})_{i,j\in\{1,\ldots,p\}})$ on the same list of trees are isomorphic if there exist automorphisms $(g_i : T_i \to T_i)_{i=1,\ldots,p}$ and $(h_i : T_i \to T_i)_{i=1,\ldots,p}$ such that $h_j \circ \phi_{ij} = \phi'_{ij} \circ g_i$ for $i, j = 1, \ldots, p$.

The two tangled chains $X$ and $X'$ are determined by the sequences $\phi_{12}, \phi_{23}, \ldots, \phi_{(p-1)p}$ and $\phi'_{12}, \phi'_{23}, \ldots, \phi'_{(p-1)p}$. The previous definition implies that $\phi_{(i-1)\ i} = h_i^{-1} \circ \phi'_{(i-1)i} \circ g_{i-1}$ and $\phi_{(i-1)i} \in A(T_i)\phi'_{(i-1)i}A(T_{i-1})$, which is again a double coset. The latter property (which is captured in the following equivalence relation) characterizes tangled chains that are isomorphic. Let $T = (T_1, T_2, \ldots, T_p)$, for two elements $(w_1, w_2, \cdots, w_{p-1})$ and $(w'_1, w'_2, \ldots, w'_{p-1})$ of $S_n^{p-1}$, we say that

$$(w_1, w_2, \ldots, w_{p-1}) \equiv_T (w'_1, w'_2, \ldots, w'_{p-1})$$

if there exist $t_i \in A(T_i)$ such that $w_i = t_i w'_i t_{i+1}$ for all $i = 1, \ldots, p-1$. Then, $\equiv_T$ is an equivalence relation and we denote by $C^T$ the set of equivalence classes modulo $\equiv_T$. Thus, the set of non-isomorphic tangled chains corresponding to the tuple $T = (T_1, T_2, \ldots, T_p)$ is in one-to-one correspondence with the elements of $C^T$. We call the elements of $C^T$ *multicosets* of $S_n$ with respect to $A(T_1) \times A(T_2) \times \cdots \times A(T_p)$. This leads us to the next theorem.

**Theorem 2.2.7** ([3]). *The number of non-isomorphic tangled chains of length p where each tree has n leaves is*

$$t(n,p) = \sum_{\lambda} \frac{\prod_{i=2}^{l(\lambda)} (2(\lambda_i + \cdots + \lambda_{l(\lambda)}) - 1)^p}{z_\lambda}, \tag{2.2.9}$$

*where the sum is over binary partitions of n.*

**Example 2.2.8.** For $n = p = 3$, the partitions of $n$ are given by $(2,1)$ and $(1,1,1)$, and the theorem gives

$$t(3,3) = \frac{1^3}{2} + \frac{3^3 \cdot 1^3}{6} = 5.$$

The first few terms of $t(n,3)$ start with

$$1, 1, 5, 151, 9944, 1196991, 226435150, 61992679960, 23198439767669,$$

see [31, A258486] for more terms.

*Proof.* The number $t(n,p)$ of non-isomorphic tangled chains with $n$ leaves and $p$ binary trees is given by

$$t(n,p) = \sum_T |C^T|,$$

where the sum is over tuples $T = (T_1, T_2, \dots, T_p)$ of binary trees with $n$ leaves and $C^T$ is the set of multicosets corresponding to $T$. Let $T = (T_1, T_2, \dots, T_p)$ be a fixed ordered list of binary trees with $n$ leaves. For $(w_1, w_2, \dots, w_{p-1}) \in S_n^{p-1}$, we denote by $C^T(w_1, w_2, \dots, w_{p-1})$ the multicoset containing $(w_1, w_2, \dots, w_{p-1})$. As in the proof of Theorem 2.2.1, we have:

$$|C^T| = \sum_{w_1 \in S_n} \sum_{w_2 \in S_n} \cdots \sum_{w_{p-1} \in S_n} \frac{1}{|C^T(w_1, w_2, \cdots, w_{p-1})|}. \tag{2.2.10}$$

We need to find a formula for $|C^T(w_1, w_2, \cdots, w_{p-1})|$ involving only the automorphism groups of the trees $T_1, T_2, \dots, T_p$. In order to do so, we define a set $A(C^T(w_1, w_2, \dots, w_{p-1}))$, which is the subgroup of all $t_1 \in A(T_1)$ such that for $i = 2, \dots, p$ there exists $t_i \in A(T_i)$ satisfying $w_i = t_i w_i t_{i+1}$. Now, suppose $t_1 \in A(C^T(w_1, w_2, \dots, w_{p-1}))$. Then there exist $t_i \in A(T_i)$ such that $w_i = t_i w_i t_{i+1}$ for $i = 2, \dots, p-1$. So, $t_i = w_i t_{i+1}^{-1} w_i^{-1}$ and, by induction, we have $t_1 = w_2 \dots w_{j-1} t_j^{-1} (w_2 \dots w_{j-1})^{-1}$ for $j = 2, \dots, p$. Thus,

$$\begin{aligned} A(C^T(w_1, w_2, \cdots, w_{p-1})) = A(T_1) \cap w_1 A(T_2) w_1^{-1} \cap \cdots \\ \cdots \cap w_1 w_2 \cdots w_{p-1} A(T_p) w_{p-1}^{-1} \cdots w_2^{-1} w_1^{-1}, \end{aligned}$$

and

$$\begin{aligned} |A(C^T(w_1, \dots, w_{k-1}))| = \sum_{i=1}^{p} \sum_{t_i \in A(T_i)} I(t_1 = w_1 t_2 w_1^{-1}) \cdot I(t_2 = w_2 t_3 w_2^{-1}) \cdots \\ \cdots I(t_{p-1} = w_{p-1} t_p w_{p-1}^{-1}). \end{aligned} \tag{2.2.11}$$

Next, we let $T' = (T_2, \ldots, T_p)$. For each $(v_2, \ldots, v_{p-1}) \in C^{T'}(w_2, \ldots, w_{p-1})$, we want to construct an element of $C^T(w_1, \ldots, w_{p-1})$. For that purpose, we can add an element $v_1 \in A(T_1)$ in the beginning of the sequence $(v_2, \ldots, v_{p-1})$ if and only if $v_1 \in A(T_1)w_1A(C^{T'}(w_2, \ldots, w_{p-1}))$. Since $A(T_1)w_1A(C^{T'}(w_2, \ldots, w_{p-1}))$ is a double coset, by Proposition 1.2.8 we have

$$\begin{aligned} |C^T(w_1, w_2, \cdots, w_{p-1})| &= \frac{|A(T_1)| \cdot |A(C^{T'}(w_2, \cdots, w_{p-1}))|}{|A(T_1) \cap w_1 A(C^T(w_2, \cdots, w_{p-1}))w_1^{-1}|} \cdot |C^T(w_2, \cdots, w_{p-1})| \\ &= \frac{|A(T_1)| \cdot |A(C^{T'}(w_2, \cdots, w_{p-1}))|}{|A(C^T(w_1, \cdots, w_{p-1}))|} \cdot |C^T(w_2, \cdots, w_{p-1})|. \end{aligned}$$

By induction on $p$ we have

$$|C^T(w_1, w_2, \cdots, w_{p-1})| = \frac{|A(T_1)| \cdot |A(T_2)| \cdots |A(T_p)|}{|A(C^T(w_1, \cdots, w_{p-1}))|}.$$

Now, Equation (2.2.10) becomes

$$\begin{aligned} |C^T| &= \sum_{w_1 \in S_n} \sum_{w_2 \in S_n} \cdots \sum_{w_{p-1} \in S_n} \frac{|A(C^T(w_1, \cdots, w_{p-1}))|}{|A(T_1)| \cdot |A(T_2)| \cdots |A(T_p)|} \\ &= \frac{\sum_{w_1 \in S_n} \sum_{w_2 \in S_n} \cdots \sum_{w_{p-1} \in S_n} |A(C^T(w_1, \cdots, w_{p-1}))|}{|A(T_1)| \cdot |A(T_2)| \cdots |A(T_p)|}. \end{aligned}$$

By (2.2.11) the numerator becomes

$$\begin{aligned} &\sum\nolimits_{(w_1, w_2, \cdots, w_{p-1})} |A(C^T(w_1, \cdots, w_{p-1}))| \\ &= \sum_{(w_1, w_2, \cdots, w_{p-1})} \sum_{T_1 \in A(T_1)} \cdots \sum_{T_p \in A(T_p)} I(t_1 = w_1 t_2 w_1^{-1}) \cdot I(t_2 = w_2 t_3 w_2^{-1}) \cdots I(t_{p-1} = w_{p-1} t_p w_{p-1}^{-1}). \end{aligned}$$

In addition, we observe that

$$I(t_1 = w_1 t_2 w_1^{-1}) \cdot I(t_2 = w_2 t_3 w_2^{-1}) \cdots I(t_{p-1} = w_{p-1} t_p w_{p-1}^{-1}) \neq 0$$

if and only if all the $t_i$ have the same type $\lambda$ (by Proposition 2.1.7). So

$$|C^T| = \sum_{\lambda} \frac{|A(T_1)_\lambda| \cdot |A(T_2)_\lambda| \cdots |A(T_p)_\lambda| \cdot z_\lambda^{p-1}}{|A(T_1)| \cdot |A(T_2)| \cdots |A(T_p)|}.$$

Hence,

$$\begin{aligned} t(p, n) &= \sum_{(T_1, \ldots, T_p)} \sum_{\lambda} \frac{|A(T_1)_\lambda| \cdot |A(T_2)_\lambda| \cdots |A(T_p)_\lambda| \cdot z_\lambda^{p-1}}{|A(T_1)| \cdot |A(T_2)| \cdots |A(T_p)|} \\ &= \sum_{\lambda} z_\lambda^{p-1} \cdot \Big( \sum_{T \in B_n} \frac{|A(T)_\lambda|}{|A(T)|} \Big)^p. \end{aligned}$$

The theorem then follows from Proposition 2.2.2. $\square$

## 2.3 Asymptotic number of non-isomorphic tanglegrams

Now that we have a formula for $t_n$, one question that we may ask is how does $t_n$ grow when $n$ tends to infinity? In order to answer this question, we first rewrite the formula for $t_n$ in the following way.

**Corollary 2.3.1** ([3])**.** The number $t_n$ of tanglegrams of size $n$ is given by

$$t_n = \frac{C_{n-1}^2 n!}{4^{n-1}} \sum_{\mu} \frac{n(n-1)\cdots(n-|\mu|+1)}{z_\mu \cdot \prod_{i=1}^{l(\mu)} \prod_{j=1}^{\mu_i - 1} (2n - 2(\mu_1 + \cdots + \mu_{i-1}) - 2j - 1)^2}, \tag{2.3.1}$$

where $C_{n-1} = \frac{1}{n}\binom{2n-2}{n-1}$ is a Catalan number (see [31, A000108] for more information), the sum is over binary partitions $\mu$ with all parts equal to a positive power of 2 and $|\mu| \leq n$ including the empty partition in which case the summand is 1.

*Proof.* Each binary partition $\lambda$ of $n$ can be rewritten as $\lambda = \mu 1^{n-|\mu|}$ where $\mu$ is a binary partition with all parts equal to a power of 2 (greater than 1). Then, $z_\lambda = z_\mu (n - |\mu|)!$ and

$$\begin{aligned}
\prod_{i=2}^{l(\lambda)} (2(\lambda_i + \cdots + \lambda_{l(\lambda)}) - 1) &= \prod_{i=1}^{l(\lambda)-1} (2(n - \lambda_1 - \cdots - \lambda_i) - 1) \\
&= \prod_{i=1}^{l(\mu)-1} (2(n - \mu_1 - \cdots - \mu_i) - 1) \cdot (2n - 2|\mu| - 1)!! \\
&= \frac{(2n-3)!!}{\prod_{i=1}^{l(\mu)} \prod_{j=1}^{\mu_i - 1} (2n - 2(\mu_1 + \cdots + \mu_{i-1}) - 2j - 1)}.
\end{aligned}$$

The fact that $(2n-3)!!/n! = C_{n-1}/2^{n-1}$ and Theorem 2.2.1 prove that Equation (2.3.1) is another way to express the formula for the number of tanglegrams of size $n$. $\square$

The first few terms of the sum corresponding to partitions $\emptyset$, 2, 4, 22 are

$$1 + \frac{n(n-1)}{2(2n-3)^2} + \frac{n(n-1)(n-2)(n-3)}{4(2n-3)^2(2n-5)^2(2n-7)^2} + \frac{n(n-1)(n-2)(n-3)}{8(2n-3)^2(2n-7)^2}.$$

More terms can be found in [21]. We use the previous expression of $t_n$ to give an asymptotic formula for the number of tanglegrams with size $n$.

**Corollary 2.3.2** ([3])**.** We have

$$\frac{t_n}{n!} = e^{\frac{1}{8}} \frac{C_{n-1}^2}{4^{n-1}} \cdot (1 + O(n^{-1})) \sim \frac{e^{\frac{1}{8}} \cdot 4^{n-1}}{\pi \cdot n^3} (1 + O(n^{-1})). \tag{2.3.2}$$

*Proof.* It suffices to estimate the sum on the right hand side of (2.3.1). First, we show that the series

$$\sum_{\mu} \frac{1}{z_\mu}$$

is convergent, where the sum is taken over all binary partitions $\mu$ that do not contain 1. To see this, we write $\mu = (2^{i_2}, 4^{i_4}, 8^{i_8}, \cdots)$ and

$$\sum_{\mu} \frac{1}{z_\mu} = \sum_{(i_2, i_4, i_8, \cdots)} \left(2^{i_2} \cdot i_2! \cdot 4^{i_4} \cdot i_4! \cdot 8^{i_8} \cdot i_8! \cdots\right)^{-1},$$

where all but finitely many of the $i_{2^j}$'s are zero in $(i_2, i_4, i_8, \cdots)$. Hence,

$$\sum_{\mu} \frac{1}{z_\mu} = \prod_{j=1}^{\infty} \left( \sum_{k=0}^{\infty} \frac{2^{-jk}}{k!} \right) = \prod_{j=1}^{\infty} e^{2^{-j}} = e.$$

Now, for each binary partition $\mu$, let

$$a_\mu = \frac{n(n-1)\cdots(n-|\mu|+1)}{\prod_{i=1}^{l(\mu)} \prod_{j=1}^{\mu_i - 1} (2n - 2(\mu_1 + \cdots + \mu_{i-1} + j) - 1)^2}.$$

Note that the numerator in $a_\mu$ can be written in the following way

$$\prod_{i=1}^{l(\mu)} (n - (\mu_1 + \cdots + \mu_{i-1})) \cdot \prod_{i=1}^{l(\mu)} \prod_{j=1}^{\mu_i - 1} (n - (\mu_1 + \cdots + \mu_{i-1} + j)).$$

For any $i \in \{1, \cdots, l(\mu)\}$ and $j \in \{1, \cdots, \mu_i - 1\}$,

$$n - (\mu_1 + \cdots + \mu_{i-1} + j) \leq 2n - 2(\mu_1 + \cdots + \mu_{i-1} + j) - 1.$$

Similarly, for any $i \in \{1, \cdots, l(\mu)\}$,

$$n - (\mu_1 + \cdots + \mu_{i-1}) \leq 2n - 2(\mu_1 + \cdots + \mu_{i-1} + 1) - 1,$$

except for the case $i = l(\mu)$, $|\mu| = n$ and $\mu_{l(\mu)} = 2$. In that case, an additional factor 2 is needed on the right hand side of last inequality. Thus, we get

$$a_\mu \leq \frac{2}{\prod_{i=1}^{l(\mu)} \prod_{j=2}^{\mu_i - 1} (2n - 2(\mu_1 + \cdots + \mu_{i-1} + j) - 1)}.$$

In particular, $a_\mu$ is at most 2. Now, assume that $\mu$ is a binary partition such that the largest term $\mu_1 \geq 4$, then

$$\begin{aligned}
\prod_{i=1}^{l(\mu)} \prod_{j=2}^{\mu_i - 1} & (2n - 2(\mu_1 + \cdots + \mu_{i-1} + j) - 1) \\
& \geq \prod_{j=2}^{\mu_1 - 1} (2n - 2j - 1) \\
& \geq (2n-5)(2n-7).
\end{aligned}$$

This shows that the contribution of the partitions $\mu$ with $\mu_1 \geq 4$ to the sum in Equation (2.3.1) is a $O(n^{-2})$. Hence,

$$\frac{t_n}{n!} = \frac{C_{n-1}^2}{4^{n-1}} \left( \sum_{\mu,\, \mu_1 \leq 2} \frac{a_\mu}{z_\mu} + O(n^{-2}) \right).$$

For $\mu = 2^k$, we have $z_\mu = 2^k k!$, and we can also show that

$$a_\mu = \prod_{j=0}^{k-1} \frac{(n-2j)(n-(2j+1))}{(2n-(4j+3))^2}.$$

Thus, we need to estimate the sum

$$\sum_{0 \le k \le n/2} \frac{1}{2^k k!} \prod_{j=0}^{k-1} \frac{(n-2j)(n-(2j+1))}{(2n-(4j+3))^2}.$$

We split this sum into two parts:

- For $k > n^{1/3}$

$$\sum_{n^{1/3} < k \le n/2} \frac{1}{2^k k!} \prod_{j=0}^{k-1} \frac{(n-2j)(n-(2j+1))}{(2n-(4j+3))^2} \le 2 \sum_{n^{1/3} < k \le n/2} \frac{1}{2^k k!} = O(2^{-n^{1/3}}).$$

- For $k \le n^{1/3}$, we have

$$\begin{aligned}
&\log\left(\prod_{j=0}^{k-1} \frac{(n-2j)(n-(2j+1))}{(2n-(4j+3))^2}\right) \\
&= -2k\log 2 + \sum_{j=0}^{k-1}\left(\log\left(1-\frac{2j}{n}\right) + \log\left(1-\frac{2j+1}{n}\right) - 2\log\left(1-\frac{4j+3}{2n}\right)\right) \\
&= -2k\log 2 + \frac{2k}{n} + O(k^3/n^2).
\end{aligned}$$

  So,

$$\prod_{j=0}^{k-1} \frac{(n-2j)(n-(2j+1))}{(2n-(4j+3))^2} = 2^{-2k}\left(1+O(k/n)\right),$$

  where the constant in the $O$-notation is independent of $k$.

Putting everything together, we have

$$\sum_\mu \frac{a_\mu}{z_\mu} = \sum_{0 \le k \le n^{1/3}} \frac{1}{2^{3k} k!} + O\left(n^{-1} \sum_{0 \le k \le n^{1/3}} \frac{k}{2^{3k} k!}\right) = e^{1/8} + O(n^{-1}),$$

and the result follows. The second part of the formula is obtained by considering the asymptotic expansion of $C_{n-1}^2$. □

We end this chapter with an asymptotic formula for the number of tangled chains of length $p > 2$, where each binary tree has $n$ leaves. We have

$$\frac{t(n,p)}{(n!)^{p-1}} = \frac{C_{n-1}^p}{2^{p(n-1)}} \sum_\mu \frac{n(n-1)\cdots(n-|\mu|+1)}{z_\mu \prod_{i=1}^{l(\mu)} \prod_{j=1}^{\mu_i - 1} (2n - 2(\mu_1 + \cdots + \mu_{i-1} + j) - 1)^p}. \tag{2.3.3}$$

Let

$$a_{\mu,p} = \frac{n(n-1)\cdots(n-|\mu|+1)}{\prod_{i=1}^{l(\mu)} \prod_{j=1}^{\mu_i - 1} (2n - 2(\mu_1 + \cdots + \mu_{i-1} + j) - 1)^p}.$$

Then,

$$a_{\mu,p} = \frac{a_\mu}{\prod_{i=1}^{l(\mu)} \prod_{j=1}^{\mu_i - 1} (2n - 2(\mu_1 + \cdots + \mu_{i-1} + j) - 1)^{p-2}},$$

where $a_\mu$ is defined in the previous proof. Now we consider several cases:

- If $\mu = 2$, then
$$a_{\mu,p} = \frac{n(n-1)}{2(2n-3)^p}.$$
- If $\mu_1 \geq 4$, then
$$a_{\mu,p} \leq \frac{2}{(2n-3)^{p-2}(2n-5)^p(2n-7)^p}.$$
- If $\mu = 2^l$, with $l \geq 2$, then
$$a_{\mu,p} \leq \frac{2}{(2n-3)^{p-2}(2n-7)^{p-2}}.$$

Putting these estimates into Equation (2.3.3), we obtain the asymptotic formula

$$\frac{t(n,p)}{(n!)^{p-1}} = \frac{C_{n-1}^p}{2^{p(n-1)}} \left( 1 + \frac{n(n-1)}{(2n-3)^p} + O(n^{-2(p-2)}) \right).$$

Note that only the empty partition contributes to the main term of $t(n,p)$.

# Chapter 3

# Random tanglegrams

It is natural to study parameters of random tanglegrams when the exact enumeration is done. Here and in the next chapter, we consider two tanglegrams that are isomorphic to be equal. Furthermore, in this chapter, we consider the uniform probability measure on the set of non-isomorphic tanglegrams on $n$ leaves. Then, as in the work of Konvalinka and Wagner in [21], we show that a typical large tanglegram looks like two independently chosen random plane binary trees. The latter fact is used to derive a number of results on the following parameters: the number of occurrences of subtrees, the distribution of root branches, the number of automorphisms and the height of a tanglegram. Cherries (a subtree of a binary tree $T$ consisting of an internal vertex with exactly two leaves as children) play important roles in the literature of tanglegrams as it is mentioned in [3]. So we will determine the expected value as well as the limiting distribution of matched cherries (two cherries whose leaves are matched to each other).

## 3.1 Comparison between the number of tanglegrams and the number of pairs of plane binary trees

In [3], the authors gave an algorithm to randomly generate a tanglegram and a number of questions were put forward. Then, in [21], Konvalinka and Wagner answered those questions using a probabilistic approach. This approach will be elaborated here.

First, let us recall the concept of the total variation distance of probability measures. Let $\pi_1$ and $\pi_2$ be two probability measures on a finite set $\Omega$. The two measures $\pi_1$ and $\pi_2$ will be defined over the same $\sigma$-algebra $\mathcal{F}$ which, for our purpose, will be the entire powerset $\mathcal{P}(\Omega)$. We have the following definition:

**Definition 3.1.1.** The total variation distance between $\pi_1$ and $\pi_2$ is the quantity

$$d(\pi_1, \pi_2) = \sup_{S \in \mathcal{F}} |\pi_1(S) - \pi_2(S)| = \sup_{S \subseteq \Omega} |\pi_1(S) - \pi_2(S)|. \tag{3.1.1}$$

**Lemma 3.1.2.** The total variation distance between $\pi_1$ and $\pi_2$ can also be rewritten as

$$d(\pi_1, \pi_2) = \frac{1}{2} \sum_{x \in \Omega} |\pi_1(x) - \pi_2(x)|. \tag{3.1.2}$$

For simplicity, we write $\pi_1(x)$ for the probability of the event $\{x\}$ given any probability measure $\pi_1$ on a finite set $\Omega$.

*Proof.* Let $A = \{x \in \Omega \mid \pi_1(x) \geq \pi_2(x)\}$, we will first show that

$$d(\pi_1, \pi_2) = \sup_{S \subseteq \Omega} |\pi_1(S) - \pi_2(S)| = \pi_1(A) - \pi_2(A).$$

Indeed, for a set $S \subseteq \Omega$, we have

$$\pi_1(S) - \pi_2(S) = \sum_{x \in S \cap A} (\pi_1(x) - \pi_2(x)) + \sum_{x \in S \setminus A} (\pi_1(x) - \pi_2(x)).$$

Since $\sum_{x \in S \cap A} (\pi_1(x) - \pi_2(x)) \geq 0$ and $\sum_{x \in S \setminus A} (\pi_1(x) - \pi_2(x)) \leq 0$, we have

$$\pi_1(S) - \pi_2(S) \leq \sum_{x \in S \cap A} (\pi_1(x) - \pi_2(x)) \leq \pi_1(A) - \pi_2(A).$$

Similarly,

$$\pi_1(S) - \pi_2(S) \geq \sum_{x \in S \setminus A} (\pi_1(x) - \pi_2(x)) \geq \pi_1(\Omega \setminus A) - \pi_2(\Omega \setminus A).$$

Now, $|\pi_1(S) - \pi_2(S)| \leq |\pi_1(A) - \pi_2(A)|$ because

$$|\pi_1(A) - \pi_2(A)| = |\pi_1(\Omega \setminus A) - \pi_2(\Omega \setminus A)|.$$

Furthermore,

$$\begin{aligned}
d(\pi_1, \pi_2) &= \sup_{S \subseteq \Omega} |\pi_1(S) - \pi_2(S)| \\
&= \pi_1(A) - \pi_2(A) \\
&= \frac{1}{2}(2\pi_1(A) - 2\pi_2(A)) \\
&= \frac{1}{2}(\pi_1(A) - \pi_2(A) + 1 - \pi_2(A) - 1 + \pi_1(A)) \\
&= \frac{1}{2}(\pi_1(A) - \pi_2(A) + \pi_2(\Omega \setminus A) - \pi_1(\Omega \setminus A) \\
&= \frac{1}{2}\Big( \sum_{x \in A} (\pi_1(x) - \pi_2(x)) + \sum_{x \in \Omega \setminus A} (\pi_2(x) - \pi_1(x)) \Big) \\
&= \frac{1}{2} \sum_{x \in \Omega} |\pi_1(x) - \pi_2(x)|.
\end{aligned}$$

$\square$

Now, a plane binary tree is a binary tree embedded in the plane so that the left child of a vertex (if it exists) is distinguishable from the right child (if it exists). It is well known that the number of plane binary trees with $n+1$ leaves is the Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$. We denote by $\mathcal{B}_n$ the set of the equivalence classes of plane binary trees with $n \geq 1$ leaves with respect

to graph isomorphisms and use the notation $\mathcal{B}_n^2$ for $\mathcal{B}_n \times \mathcal{B}_n$. The set $\mathcal{B}_n$ is just the set of non-plane binary trees with $n$ leaves which are enumerated by the Wedderburn-Etherington numbers with an alternative formula given in Theorem 2.2.3.
Next, we define two probability measures on $\mathcal{B}_n^2$. The first probability, denoted by $\nu_n^{(T)}$, is the probability induced by the uniform probability on random tanglegrams (the two components simply being the two halves of a tanglegram). For an element $(B_1, B_2) \in \mathcal{B}_n^2$, this probability is defined by

$$\nu_n^{(T)}(B_1, B_2) = \frac{1}{t_n} \sum_{\lambda} z_\lambda \frac{|A(B_1)_\lambda|}{|A(B_1)|} \frac{|A(B_2)_\lambda|}{|A(B_2)|}. \tag{3.1.3}$$

Here,

$$\sum_{\lambda} z_\lambda \frac{|A(B_1)_\lambda|}{|A(B_1)|} \frac{|A(B_2)_\lambda|}{|A(B_2)|}$$

is the number of non-isomorphic tanglegrams which have $B_1$ and $B_2$ as trees (see [Theorem 2.2.1, Proof]).
The second probability, denoted by $\nu_n^{(P)}$, is the probability obtained by choosing two plane binary trees on $n$ leaves uniformly and independently at random. The probability that a pair of randomly chosen plane binary tree is isomorphic to $(B_1, B_2)$ is given by

$$\nu_n^{(P)}(B_1, B_2) = \left(\frac{2^{n-1}}{C_{n-1}}\right)^2 \cdot \frac{1}{|A(B_1)||A(B_2)|}. \tag{3.1.4}$$

Here, we note that the set of rotations (by rotation we mean the action of permuting the two branches) of the $n-1$ internal vertices of a plane binary tree generates a group of order $2^{n-1}$. This group acts on the set of possible representations of a plane binary tree. Then, the stabilizers have cardinality $|A(B_i)|$ $(i = 1, 2)$ and by the orbit-stabilizer theorem, the number of orbits which is the number of distinct representations, is given by $2^{n-1}/|A(B_i)|$. We have the following theorem:

**Theorem 3.1.3.** *The total variation distance $d(\nu_n^{(T)}, \nu_n^{(P)})$ goes to 0 as $n \to \infty$, specifically $d(\nu_n^{(T)}, \nu_n^{(P)}) = O(n^{-1/2})$. Moreover, there exist positive constants $M_1$ and $M_2$ such that we have:*

$$\nu_n^{(T)}(S) \leq M_1 \nu_n^{(P)}(S) + O(n^{-1}) \quad \text{and} \quad \nu_n^{(P)}(S) \leq M_2 \nu_n^{(T)}(S) + O(n^{-1}), \tag{3.1.5}$$

*for every subset $S$ of $\mathcal{B}_n^2$.*

We will prove this theorem using two lemmas. The first lemma provides information on the size of $|A(T)_\lambda|$ for a binary partition $\lambda$ and a binary tree $|A(T)|$. This lemma uses a specific subtree of $T$ called *cherry* to bound $|A(T)_\lambda|$. Formally, a cherry of a binary tree $T$ is a subtree of $T$ consisting of an internal vertex with exactly two leaves as children. Then, the first lemma is stated as follows:

**Lemma 3.1.4** ([21]). Let $\mu(s) = 2^s 1^{n-2s}$ be the partition of $n$ consisting of $s$ twos and $n - 2s$ ones, and let $c(T)$ denote the number of cherries of a binary tree $T$. We have the inequalities:

$$\binom{c(T)}{s} \leq |A(T)_{\mu(s)}| \leq \binom{c(T)+s-1}{s}.$$

*Proof.* We can choose $s$ cherries from the $c(T)$ cherries of $T$ and obtain one automorphism of type $\mu(s) = 2^s 1^{n-2s}$ by permuting the leaves in each of the $s$ cherries and fixing the remaining leaves. This proves the inequality

$$\binom{c(T)}{s} \leq |A(T)_{\mu(s)}|.$$

For the inequality

$$|A(T)_{\mu(s)}| \leq \binom{c(T)+s-1}{s}, \tag{3.1.6}$$

we define a polynomial

$$P(T,u) = \sum_{s\geq 0} |A(T)_{\mu(s)}| \cdot u^s.$$

Let $T_1$ and $T_2$ be the two branches of $T$. One can easily see that if $T_1 \neq T_2$, then $A(T) = A(T_1) \times A(T_2)$ (see Corollary 1.1.7) so $P(T,u) = P(T_1,u) \cdot P(T_2,u)$. Otherwise, if $T_1 = T_2$ then

$$P(T,u) = P(T_1,u)^2 + |A(T_1)| \cdot u^{|T_1|}. \tag{3.1.7}$$

In Equation (3.1.7), the term $P(T_1,u)^2$ accounts for the permutations of type $\mu(s)$ that does not exchange the two branches $T_1$ and $T_2$ of $T$. On the otherside, the term $|A(T_1)| \cdot u^{|T_1|}$ accounts for the permutations of type $\mu(s)$ that exchanges the two branches $T_1$ and $T_2$ of $T$. Indeed, the number of automorphisms which exchanges the two trees $T_1$ and $T_2$ is given by $|A(T_1)||A(T_1)_{\lambda/2}|$ for a binary partition $\lambda$ with all part at least 2 (see Proposition 2.1.8). If $\lambda = 2^s = (2,2,\ldots,2)$ then $\lambda/2 = (1,1,\ldots,1)$ and we only have one element in $A(T_1)_{\lambda/2}$ which is the identity, so $|A(T_1)_{\lambda/2}| = 1$. In addition, if a permutation exchanges the two trees and have type $\lambda = 2^s$ then $s$ must be equal to $|T_1|$.

Next, we want to show that the coefficient of $u^s$ in $P(T,u)$ is always less than or equal to $\binom{c(T)+s-1}{s}$, which is the coefficient of $u^s$ in $(1-u)^{-c(T)}$. Denoting by $\preceq$ the coefficient-wise inequality of polynomials or power series, we will prove by induction on $c(T)$ that

$$P(T,u) \preceq (1-u)^{-c(T)}.$$

We first notice that the degree of $P(T,u)$ is at most $|T|/2$. The inequality is true for the one leaf tree or the tree with only one cherry. Now, we proceed with the induction.

If $T_1 \neq T_2$ then

$$P(T,u) = P(T_1,u)P(T_2,u) \preceq (1-u)^{-c(T_1)}(1-u)^{-c(T_2)} = (1-u)^{-c(T)},$$

and we are done. If the two branches are the same, we need to be more careful since

$$P(T,u) = P(T_1,u)^2 + |A(T_1)| \cdot u^{|T_1|}.$$

The inequality follows in the same way, except perhaps for the coefficient of $u^{|T_1|}$ which we need to compute directly. For that purpose, we first set $t = |T_1|$. Since the degree of $P(T_1, u)$ is at most $t/2$, the contribution of $P(T_1, u)^2$ to the coefficient of $u^t$ is the square of the coefficient of $u^{t/2}$ in $P(T_1, u)$ (if such a coefficient exists ). Hence, the coefficient of $u^t$ is at most

$$\binom{c(T_1) + t/2 - 1}{t/2}^2 + |A(T_1)|,$$

where we interpret the binomial coefficient as 0 if $t$ is odd. Now, we show that $|A(T_1)| \leq 2^{2c(T_1)-1}$ by induction. This is true for $c(T_1) = 1$. For the induction step, we let $S_1$ and $S_2$ be the branches of $T_1$. So

- if $S_1 \neq S_2$, then

$$|A(T_1)| = |A(S_1)||A(S_2)| \leq 2^{2c(S_1)-1} \cdot 2^{2c(S_2)-1} = 2^{2c(T_1)-2} < 2^{2c(T_1)-1}.$$

- Otherwise, $|A(T_1)| = 2|A(S_1)|^2 \leq 2 \cdot (2^{2c(S_1)-1})^2 = 2^{2c(T_1)-1}$,

which concludes the proof of the inequality $|A(T_1)| \leq 2^{2c(T_1)-1}$. Hence, we have

$$\binom{c(T_1) + t/2 - 1}{t/2}^2 + |A(T_1)| \leq \binom{c(T_1) + t/2 - 1}{t/2}^2 + 2^{2c(T_1)-1},$$

and what is remaining to show is that this is less than or equal to $\binom{c(T)+t-1}{t}$. We first note that $c(T_1) \leq t/2$. The quantity $\binom{c(T)+t-1}{t} = \binom{2c(T_1)+t-1}{t}$ counts the number of ways to select $t$ elements from the set $\{1, 2, \ldots, 2c(T_1)\}$, repetitions allowed. The expression $\binom{c(T_1)+t/2-1}{t/2}^2$ counts those choices where the same number of elements is taken from the two halves $\{1, 2, \ldots, c(T_1)\}$ and $\{c(T_1) + 1, \ldots, 2c(T_1)\}$ of the set (zero if $t$ is odd). For each subset of $\{1, \ldots, 2c(T_1) - 1\}$, of which there are exactly $2^{2c(T_1)-1}$, we create a selection where the numbers in the two halves are not the same: we add an appropriate number of copies (i.e., as many as needed to create a multiset of $t$ elements) of the element $2c(T_1)$ unless it creates a balanced selection, in which case we increase the number of the least element appropriately instead. This creates an injection that proves the inequality and hence completes the induction proof of (3.1.6). □

For a given positive integer $n$, we denote by $R$ the set of partitions $\lambda$ of $n$ of the form $\lambda = 2^s 1^{n-2s}$. As in the proof of Corollary 2.3.2, we will see that only these partitions really matters. The estimate we found for $|A(T)_\lambda|$ ($\lambda \in R$) in the previous lemma will be useful to establish the second lemma which states:

**Lemma 3.1.5** ([21]). (1) There exists an absolute constant $K$ such that

$$\frac{1}{n!} \sum_{\lambda \in R} z_\lambda |A(B_1)_\lambda||A(B_2)_\lambda| \leq K,$$

for all possible pairs $(B_1, B_2)$ of binary trees with $n$ leaves.

(2) If we further assume that $c(B_1), c(B_2) \geq \alpha n$ for some fixed constant $\alpha$, then

$$\frac{1}{n!}\sum_{\lambda\in R} z_\lambda |A(B_1)_\lambda||A(B_2)_\lambda| = \exp\left(\frac{2c(B_1)c(B_2)}{n^2}\right) + O(n^{-1}).$$

where the constant implied by the $O$-term only depends on $\alpha$.

*Proof.* From the previous lemma we have

$$\begin{aligned}\frac{1}{n!}\sum_{\lambda\in R} z_\lambda |A(B_1)_\lambda||A(B_2)_\lambda| &\leq \frac{1}{n!}\sum_{0\leq s\leq n/2} z_{\mu(s)} \binom{c(B_1)+s-1}{s}\binom{c(B_2)+s-1}{s} \\ &\leq \frac{1}{n!}\sum_{0\leq s\leq n/2} z_{\mu(s)} \binom{n/2+s-1}{s}^2. \end{aligned} \tag{3.1.8}$$

By definition we have $z_{\mu(s)} = 2^s s!(n-2s)!$, so it remains to prove that the quantity

$$\frac{1}{n!}\sum_{0\leq s\leq n/2} 2^s s!(n-2s)!\binom{n/2+s-1}{s}^2$$

is bounded by a constant.
In order to do so, we make use of the inequality

$$(n-2s)! = n!\prod_{j=0}^{2s-1}(n-j)^{-1} \leq n!(n-2s)^{-2s} \tag{3.1.9}$$

and split the sum on the right hand side of (3.1.8) into three parts: the first part is

$$\begin{aligned}\frac{1}{n!}\sum_{0\leq s\leq \sqrt{n}} 2^s s!(n-2s)!\binom{n/2+s-1}{s}^2 &\leq \sum_{0\leq s\leq \sqrt{n}} 2^s s! \frac{(n-2s)!}{n!}\frac{(n/2+s-1)^{2s}}{(s!)^2} \\ &\leq \sum_{0\leq s\leq \sqrt{n}} 2^s (n-2s)^{-2s}\frac{(n/2+s-1)^{2s}}{s!} \\ &\leq \sum_{0\leq s\leq \sqrt{n}} 2^s (n-2\sqrt{n})^{-2s}\frac{(n/2+\sqrt{n})^{2s}}{s!} \\ &= \sum_{0\leq s\leq \sqrt{n}} \frac{2^{-s}}{s!}\left(\frac{n+2\sqrt{n}}{n-2\sqrt{n}}\right)^{2s} \\ &\leq \sum_{s\geq 0} \frac{2^{-s}}{s!}\left(\frac{\sqrt{n}+2}{\sqrt{n}-2}\right)^{2s} = \exp\left(\frac{(\sqrt{n}+2)^2}{2(\sqrt{n}-2)^2}\right),\end{aligned}$$

which converges to $e^{1/2}$ when $n\to\infty$, so it is bounded.
For the second part, we have to compute the sum over $\sqrt{n} < s \leq \frac{n}{6}$. Since $n-2s > n/2+s-1$ for $s \leq n/6$, we have

$$\frac{1}{n!}\sum_{\sqrt{n}<s\leq n/6} 2^s s!(n-2s)!\binom{n/2+s-1}{s}^2 \leq \sum_{\sqrt{n}<s\leq n/6} 2^s(n-2s)^{-2s}\frac{(n/2+s-1)^{2s}}{s!}$$
$$\leq \sum_{\sqrt{n}<s\leq n/6} \frac{2^s}{s!} = O\left(\frac{2^{\lceil\sqrt{n}\rceil}}{\lceil\sqrt{n}\rceil!}\right), \tag{3.1.10}$$

which converges to zero as $n \to \infty$, therefore it is also bounded. Lastly,

$$\frac{1}{n!}\sum_{n/6<s\leq n/2} 2^s s!(n-2s)!\binom{n/2+s-1}{s}^2 \leq \sum_{n/6<s\leq n/2} \frac{2^s}{\binom{n-s}{s}}\frac{(n-s)!}{n!}(2^{n/2+s-1})^2.$$

Now, modifying Equation (3.1.9) a little bit gives

$$(n-s)! = n!\prod_{j=0}^{s-1}(n-j)^{-1} \leq n!(n-s)^{-s}.$$

Thus,

$$\frac{1}{n!}\sum_{n/6<s\leq n/2} 2^s s!(n-2s)!\binom{n/2+s-1}{s}^2 \leq \sum_{n/6<s\leq n/2} 2^{n/2}(n/2)^{-s}2^{2n}$$
$$\leq (n/2)\cdot 2^{5n/2}(n/2)^{-n/6}, \tag{3.1.11}$$

which also goes to 0 as $n \to \infty$, so it is bounded. This completes the proof of the first part of Lemma 3.1.5.
For the second part of the lemma, we suppose that $c(B_1), c(B_2) \geq \alpha n$. Now the inequality

$$\binom{c(B_i)}{s} \leq |A(B_i)_{\mu(s)}| \leq \binom{c(B_i)+s-1}{s}$$

implies

$$|A(B_i)_{\mu(s)}| \leq \frac{(c(B_i)+s-1)^s}{s!} \leq \frac{c(B_i)^s}{s!}\left(1+\frac{s}{c(B_i)}\right)^s.$$

Hence, for $s \leq \sqrt{n}$, we have

$$|A(B_i)_{\mu(s)}| = \frac{c(B_i)^s}{s!}\Big(1+O(s^2/n)\Big).$$

We know from the two estimates (3.1.10) and (3.1.11) that the partition $\mu(s)$ with $s > \sqrt{n}$ only contribute an error less than $O(n^{-1})$ to the sum

$$\frac{1}{n!}\sum_{\lambda\in R} z_\lambda|A(B_1)_\lambda||A(B_2)_\lambda|.$$

Thus, we can focus on the values of $s$ with $s \leq \sqrt{n}$. For those, using the estimate (3.1.9), we have

$$(n-2s)! = n! \cdot n^{-2s}(1+O(s^2/n)).$$

Putting everything together yields

$$\begin{aligned}
\frac{1}{n!}\sum_{\lambda\in R} z_\lambda |A(B_1)_\lambda||A(B_2)_\lambda| &= \frac{1}{n!}\sum_{0\le s\le\sqrt{n}} 2^s s!(n-2s)!|A(B_1)_{\mu(s)}||A(B_2)_{\mu(s)}| + O(n^{-1}) \\
&= \sum_{0\le s\le\sqrt{n}} 2^s s! n^{-2s}\frac{c(B_1)^s}{s!}\frac{c(B_2)^s}{s!}(1+O(s^2/n)) + O(n^{-1}) \\
&= \sum_{s\ge 0}\frac{1}{s!}\Big(\frac{2c(B_1)c(B_2)}{n^2}\Big)^s - \sum_{s>\sqrt{n}}\frac{1}{s!}\Big(\frac{2c(B_1)c(B_2)}{n^2}\Big)^s \\
&+ O\Big(n^{-1}\sum_{s\ge 0}\frac{s^2}{s!}\Big(\frac{2c(B_1)c(B_2)}{n^2}\Big)^s\Big) + O(n^{-1}).
\end{aligned}$$

Since $c(B_1), c(B_2) \le \frac{n}{2}$, we have $2c(B_1)c(B_2)/n^2 \le 1/2$, which implies that the infinite sums $\sum_{s>\sqrt{n}}\frac{1}{s!}\Big(\frac{2c(B_1)c(B_2)}{n^2}\Big)^s = O(n^{-1})$ and $\sum_{s\ge 0}\frac{s^2}{s!}\Big(\frac{2c(B_1)c(B_2)}{n^2}\Big)^s$ is bounded. Hence,

$$\frac{1}{n!}\sum_{\lambda\in R} z_\lambda |A(B_1)_\lambda||A(B_2)_\lambda| = \exp(\frac{2c(B_1)c(B_2)}{n^2}) + O(n^{-1}).$$

This completes the proof of Lemma 3.1.5. □

Now, we can proceed to the proof of the main theorem of this section:

*Proof of Theorem 3.1.3.* We first rewrite the total variation distance between $\nu_n^{(P)}$ and $\nu_n^{(T)}$ using Lemma 3.1.2 and obtain:

$$d(\nu_n^{(T)}, \nu_n^{(P)}) = \frac{1}{2}\sum_{(B_1,B_2)\in\mathcal{B}_n^2} |\nu_n^{(T)}(B_1,B_2) - \nu_n^{(P)}(B_1,B_2)|.$$

Next, we have

$$\nu_n^{(T)}(B_1,B_2) = \frac{1}{t_n}\sum_\lambda z_\lambda \frac{|A(B_1)_\lambda|}{|A(B_1)|}\frac{|A(B_2)_\lambda|}{|A(B_2)|}.$$

We observe that only partitions $\lambda$ belonging to $R$, i.e. of the form $\lambda = 2^s 1^{n-2s}$ really matter here. More precisely,

$$\sum_{(B_1,B_2)\in\mathcal{B}_n^2}\frac{1}{t_n}\sum_{\lambda\notin R} z_\lambda \frac{|A(B_1)_\lambda|}{|A(B_1)|}\frac{|A(B_2)_\lambda|}{|A(B_2)|} = O(n^{-1}).$$

Indeed, from Proposition 2.2.2 we have

$$\begin{aligned}
\sum_{(B_1,B_2)}\frac{1}{t_n}\sum_{\lambda\notin R} z_\lambda \frac{|A(B_1)_\lambda|}{|A(B_1)|}\frac{|A(B_2)_\lambda|}{|A(B_2)|} &= \frac{1}{t_n}\sum_{\lambda\notin R} z_\lambda\Big(\sum_{B_1}\frac{|A(B_1)_\lambda|}{|A(B_1)|}\Big)^2 \\
&= \frac{1}{t_n}\sum_{\lambda\notin R}\frac{\prod_{i=2}^{l(\lambda)}(2(\lambda_i+\cdots+\lambda_{l(\lambda)})-1)^2}{z_\lambda}.
\end{aligned}$$

Using the same reasoning as in the proof of Corollary 2.3.2, we get

$$\frac{1}{t_n}\sum_{\lambda\notin R}\frac{\prod_{i=2}^{l(\lambda)}(2(\lambda_i+\cdots+\lambda_{l(\lambda)})-1)^2}{z_\lambda}=O(n^{-1}).$$

Thus we can restrict ourselves to summations over elements of $R$ in the following:

$$d(\nu_n^{(T)},\nu_n^{(P)})=\frac{1}{2}\sum_{(B_1,B_2)\in\mathcal{B}_n^2}\left|\frac{1}{t_n}\sum_{\lambda\in R}z_\lambda\frac{|A(B_1)_\lambda|}{|A(B_1)|}\frac{|A(B_2)_\lambda|}{|A(B_2)|}-\nu_n^{(P)}(B_1,B_2)\right|+O(n^{-1}).$$

Using the asymptotic formula (2.3.2) for $t_n$ established in Chapter 2 and (3.1.4), we obtain:

$$\frac{1}{t_n|A(B_1)||A(B_2)|}=\nu_n^{(P)}(B_1,B_2)\cdot\frac{1}{e^{1/8}n!}(1+O(n^{-1})),$$

for any given pair $(B_1,B_2)$, which gives us:

$$d(\nu_n^{(T)},\nu_n^{(P)})=\frac{1}{2}\sum_{(B_1,B_2)\in\mathcal{B}_n^2}\nu_n^{(P)}(B_1,B_2)\left|\frac{1}{e^{1/8}n!}\sum_{\lambda\in R}z_\lambda|A(B_1)_\lambda||A(B_2)_\lambda|-1\right|+O(n^{-1}).\quad(3.1.12)$$

Furthermore, using an approach similar to [27] or [15, Examples III.14 and IX.25], we find that the number of cherries in a random plane binary tree with $n$ leaves asymptotically follows a normal distribution, with mean $n(n-1)/(4n-6)\sim n/4$ and variance $O(n)$. Hence, by Chebyshev's inequality, $|c(B)-n/4|>p$ occurs with probability of order $O(n/p^2)$, where $B$ is a random plane binary tree with $n$ leaves. Setting $p=n/8$, we find that $c(B)<n/8$ only holds with probability $O(n^{-1})$ when a plane binary tree $B$ is selected uniformly at random. Hence, by the first part of Lemma 3.1.5, if $c(B_1)<n/8$ or $c(B_2)<n/8$ then the sum

$$\frac{1}{n!}\sum_{\lambda\in R}z_\lambda|A(B_1)_\lambda||A(B_2)_\lambda|$$

only contributes $O(n^{-1})$ to the quantity (3.1.12). Otherwise, we use the second part of Lemma 3.1.5 to obtain

$$d(\nu_n^{(T)},\nu_n^{(P)})=\sum_{(B_1,B_2)\in\mathcal{B}_n^2}\nu_n^{(P)}(B_1,B_2)\left|\exp\left(\frac{2c(B_1)c(B_2)}{n^2}-\frac{1}{8}\right)-1\right|+O(n^{-1}).$$

The sum can be expressed as the expected value of

$$\left|\exp\left(\frac{2c(B_1)c(B_2)}{n^2}-\frac{1}{8}\right)-1\right|$$

with respect to the probability $\nu_n^{(P)}$. Around $(x_1,x_2)=(1/4,1/4)$ the Taylor expansion of the function $\exp\left(2x_1x_2-\frac{1}{8}\right)-1$ gives

$$\left|\exp\left(2x_1x_2-\frac{1}{8}\right)-1\right|=O\Big(\left|x_1-\frac{1}{4}\right|+\left|x_2-\frac{1}{4}\right|\Big),$$

which also holds for bounded $x_1$ and $x_2$. Here, $x_1=c(B_1)/2$ and $x_2=c(B_2)/2$ are bounded so the total variation distance is estimated by

$$d(\nu_n^{(T)}, \nu_n^{(P)}) = O\Big(\mathbb{E}_n^{(P)}\Big(\Big|\frac{c(B_1)}{n} - \frac{1}{4}\Big| + \Big|\frac{c(B_2)}{n} - \frac{1}{4}\Big|\Big) + n^{-1}\Big),$$

where $\mathbb{E}_n^{(P)}$ is the expected value with respect to $\nu_n^{(P)}$. We let $\mathbb{E}_n^{(B)}$ denote the expected value when a plane binary tree $B$ is chosen uniformly at random. Since $B_1$ and $B_2$ are independent, we have

$$\mathbb{E}_n^{(P)}\Big(\Big|\frac{c(B_1)}{n} - \frac{1}{4}\Big| + \Big|\frac{c(B_2)}{n} - \frac{1}{4}\Big|\Big) = 2\mathbb{E}_n^{(B)}\Big(\Big|\frac{c(B)}{n} - \frac{1}{4}\Big|\Big).$$

Since the function $\phi(x) = x^2$ is convex for $x \in \mathbb{R}$, Jensen's inequality gives us:

$$\mathbb{E}_n^{(B)}\Big(\Big|\frac{c(B)}{n} - \frac{1}{4}\Big|\Big) \leq \Big(\mathbb{E}_n^{(B)}\Big(\Big(\frac{c(B)}{n} - \frac{1}{4}\Big)^2\Big)\Big)^{1/2} = O(n^{-1/2}),$$

by the fact that the variance of $c(B)/n$ is of order $O(n^{-1})$. Putting everything together, we get

$$d(\nu_n^{(T)}, \nu_n^{(P)}) = O(n^{-1/2}),$$

which is the first part of Theorem 3.1.3. Now, we prove the second part of Theorem 3.1.3. For a subset $S \subseteq B_n^2$, we have:

$$\begin{aligned}\nu_n^{(T)}(S) = \sum_{(B_1,B_2)\in S} \nu_n^{(T)}(B_1, B_2) &= \sum_{(B_1,B_2)\in S} \frac{1}{t_n} \sum_{\lambda} z_\lambda \frac{|A(B_1)_\lambda||A(B_2)_\lambda|}{|A(B_1)||A(B_2)|} \\ &= \sum_{(B_1,B_2)\in S} \frac{1}{t_n} \sum_{\lambda\in R} z_\lambda \frac{|A(B_1)_\lambda||A(B_2)_\lambda|}{|A(B_1)||A(B_2)|} + O(n^{-1})\end{aligned}$$

since

$$\frac{1}{t_n} \sum_{\lambda\notin R} z_\lambda \frac{|A(B_1)_\lambda||A(B_2)_\lambda|}{|A(B_1)||A(B_2)|} = O(n^{-1}).$$

Using the fact that

$$\frac{1}{t_n|A(B_1)||A(B_2)|} = \nu_n^{(P)}(B_1, B_2) \cdot \frac{1}{e^{1/8}n!} \cdot (1 + O(n^{-1})),$$

we have

$$\nu_n^{(T)}(S) = \sum_{(B_1,B_2)\in S} \nu_n^{(P)}(B_1, B_2) \cdot \frac{1}{e^{1/8}n!} \cdot \sum_{\lambda\in R} |A(B_1)_\lambda||A(B_2)_\lambda| + O(n^{-1}).$$

From the first part of Lemma 3.1.5, we have

$$\nu_n^{(T)}(S) \leq \sum_{(B_1,B_2)\in S} \nu_n^{(P)}(B_1, B_2) \frac{K}{e^{1/8}} + O(n^{-1}) = \frac{K}{e^{1/8}} \nu_n^{(P)}(S) + O(n^{-1}).$$

The second inequality is obtained by noticing that

$$\frac{1}{e^{1/8}n!} \cdot \sum_{\lambda\in R} |A(B_1)_\lambda||A(B_2)_\lambda| \geq \frac{1}{e^{1/8}},$$

which we can get by taking the partition $\lambda = \mu(0) = 1^n$ consisting solely of ones into account. Indeed, for this case, we have $z_\lambda = n!$ and $|A(B_1)_\lambda| = |A(B_2)_\lambda| = 1$. This completes the proof of Theorem 3.1.3.

□

## 3.2 Distribution of different parameters of a tanglegram

The main theorem of this chapter tells us that the measures $v_n^{(T)}$ and $v_n^{(P)}$ are almost the same. It follows that the behavior of various parameters of a tanglegram can be obtained directly from plane binary trees. Cherries play an important role in the literature of tanglegrams. So, the first corollary of the main theorem gives an estimation of the number of cherries in one tree (top or bottom) of a tanglegram. More generally, the next corollary deals with the number of copies of a fixed rooted binary tree $B$ occurring as a fringe subtree (i.e. a subtree consisting of a vertex and all its successors) in one half of a tanglegram. Furthermore, we remark that knowledge about generating functions and symbolic methods are required starting from here. We refer to [15] for general background on generating functions and symbolic methods. We recall that the size of a tanglegram is the number of leaves. The size of a rooted binary tree $B$ will be its number of leaves and we denote by $|B|$ this size ($|T|$ for a tanglegram $T$). Then, our first corollary is stated as follows:

**Corollary 3.2.1** ([21])**.** The average number of cherries in the top (or bottom) tree of a random tanglegram of size $n$ is asymptotically $n/4$; generally, the average number of occurrences of a fixed binary tree $B$ is asymptotically equal to $\mu_B n$, where the constant $\mu_B$ is given by $2^{1-|B|}/|A(B)|$. Moreover, the number of occurrences is asymptotically normally distributed: if $X_{n,B}$ denotes the number of occurrences of $B$ in the top half of a random tanglegram of size $n$, then we have:

$$\lim_{n\to\infty} \nu_n^{(T)}(X_{n,B} \le \mu_B n + x\sigma_B\sqrt{n}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt,$$

for every real $x$, where the constant $\sigma_B$ is defined as

$$\sigma_B^2 = 2^{1-|B|}/|A(B)| + 4^{1-|B|}(1-2|B|)/|A(B)|^2.$$

In particular, for cherries $C$ we have $\sigma_C = \frac{1}{4}$.

*Proof.* The corollary is a consequence of a similar statement for plane binary trees which we will proceed to prove. First, the number of plane binary trees isomorphic to a binary tree $B$ is $2^{|B|-1}/|A(B)|$. Thus, the bivariate generating function $Y(x,u)$ for plane binary trees, where the exponent of $x$ marks the number of leaves and $u$ the number of occurrences of $B$, satisfies the following functional equation

$$Y(x,u) = x + Y(x,u)^2 + (u-1)\frac{2^{|B|-1}}{|A(B)|}x^{|B|}.$$

The term $x$ corresponds to the tree with only one leaf. The second term $Y(x,u)^2$ corresponds to the fact that the number of occurrences of $B$ in a tree $T$ can be obtained by the sum of the occurrences of $B$ in the two branches. However, if $T$ is isomorphic to $B$ then there is no occurrences of $B$ in the two branches of $T$. But, since the tree $T$ itself is an occurrence of $B$, then the number of occurrences of $B$ in $T$ is one in this case. This fact is taken into account in the last term. The functional equation can be solved explicitly: using the abbreviations $a = 2^{|B|-1}/|A(B)|$ and $b = |B|$ we have

$$Y(x,u) = \frac{1}{2}\left(1 - \sqrt{1 - 4x + 4a(1-u)x^b}\right). \tag{3.2.1}$$

In order to study this functional equation, we use the approach given in Example IX.26 of [15] which uses singularity analysis. The function $Y(x,u)$ has a dominant singularity of square root type at the smallest zero of the polynomial $1 - 4x + 4a(1-u)x^b$. Moreover, the explicit form of the generating function studied in Example IX.26 of [15] only differs from the generating $Y(x,u)$ by a factor $x$ in the denominator and the factor $a$ in front of $(1-u)$. Hence, all the conditions needed in Example IX.26 of [15] are still satisfied. The computations are then carried out in the same way. Thus as in Proposition IX.16 of [15], using the quasi-power theorem ([15, Theorem IX.12]), the number of occurrences of a binary tree $B$ in a random binary tree follows a Gaussian limit distribution, with mean $\mu_n$ and variance $\sigma_n$ that satisfy

$$\mu_n \sim \frac{n}{4^{b-1}} \cdot a = n \cdot \frac{2^{1-b}}{|A(B)|}$$

and

$$\sigma_n^2 \sim n \cdot \left(a\frac{1}{4^{b-1}} + a^2\frac{1-2b}{4^{2(b-1)}}\right) = n \cdot \left(\frac{2^{1-b}}{|A(B)|} + 4^{1-b}\frac{(1-2b)}{|A(B)|^2}\right).$$

Another way to prove the statement for plane binary tree is to use the approach given in [10, Section 3.3]. Now, by Theorem 3.1.3, the probability of the event $X_{n,B} \leq \mu_B n + x\sigma_B\sqrt{n}$ can only change by an error of order $O(n^{-1/2})$ so the central limit theorem follows for tanglegrams. Furthermore, the average value is of order $O(n)$ so to go from plane binary trees to tanglegrams, we get a change of at most $O(n \cdot d(\nu_n^{(T)}, \nu_n^{(P)})) = O(n^{-1/2})$, which does not affect the main term. □

An analogous statement provides knowledge about the root branches:

**Corollary 3.2.2** ([21])**.** The limiting probability that one of the root branches of the top (or bottom) tree of a random tanglegram consists of a single leaf is $1/2$. Generally, the limiting probability where a fixed binary tree $B$ occurs as one of the two root branches of the top tree of a random tanglegram is $2^{-|B|}/|A(B)|$.

*Proof.* As in the first corollary, the result comes from the analogous result for plane binary trees. Again, the number of plane binary trees isomorphic to $B$ is given by $2^{|B|-1}/|A(B)|$. Indeed, as in the definition of $\nu_n^{(P)}$, the group generated by the rotations of the $|B|-1$ internal vertices of $B$ acts on the different plane representations of $B$. The stabilizer is exactly the automorphism group of $B$ and the number of orbits, which is the number of different plane representations is given by $2^{|B|-1}/|A(B)|$ by the orbit-stabilizer theorem. So for $n \geq 2|B|$, the number of binary trees for which one of the two branches is isomorphic to $B$ is $2 \cdot \frac{2^{|B|-1}}{|A(B)|} \cdot C_{n-|B|}$. The latter statement comes from the fact that one of the branches is isomorphic to $B$ and the other one is just a plane binary tree with $n - |B|$ leaves. Dividing the quantity $2 \cdot \frac{2^{|B|-1}}{|A(B)|} \cdot C_{n-|B|}$ by the total number of plane binary trees with $n$ leaves $C_{n-1}$ and taking the limit as $n \to \infty$ gives the result for plane binary trees. Then, the result for tanglegrams follows from Theorem 3.1.3. □

The following corollary is about the height i.e. the length of the longest path from the root to a leaf. Following the same idea as before, the result is carried over from plane binary trees.

**Corollary 3.2.3** ([21])**.** The average height of the top (or bottom) tree of a random tanglegram is asymptotically equal to $2\sqrt{\pi n}$ and the height asymptotically follows the theta distribution: if $H_n$ denotes the height of the top half of a random tanglegram of size $n$, then we have:

$$\lim_{n\to\infty} \nu_n^{(T)}(H_n \geq x\sqrt{n}) = \Theta(x) = \sum_{j\geq 1} e^{-j^2x^2}(4j^2x^2 - 2),$$

for every positive real number $x$.

We only give a brief sketch of the proof omitting some details.

*Skecth of the proof.* The limit theorem is again carried over from the analogous statement for plane binary trees (see, [15, Proposition VII.16] ). However, we need to be more careful when we deal with the average height because trees with height of linear order might generate an error term of order $\sqrt{n}$. In order to deal with that case, we look at the probability for a random plane binary tree to have a height greater than $h$ for a given $h \geq 0$. Let $\mathcal{B}_n^{[h]}$ be the set of plane binary trees with $n+1$ leaves whose heights are less than or equal to $h$ ($n \geq 1$). Then, by ([13, Theorem 1.3]), the number of plane binary trees whose height is greater than $h$ is given by

$$|\mathcal{B}_n| - |\mathcal{B}_n^{[h]}| = O(|\mathcal{B}_n| n^{3/2} e^{-h^2/(4n)}).$$

Thus, the probability for the height of a random plane binary tree with $n+1$ leaves to be greater than $h$ is $O(n^{3/2}e^{-h^2/(4n)})$. Next, we set $h = n^{2/3}$ and apply the second part of Theorem 3.1.3. It follows that the probability for the height of one half of a random tanglegram to be greater than $n^{2/3}$ is of order at most $O(n^{-1})$. Hence, trees with height greater than $n^{2/3}$ only contribute at most $O(1)$ to the average height. Now, we consider trees with height less than or equal to $n^{2/3}$. We apply the first part of Theorem 3.1.3 to see that the average height only changes by at most $O(n^{2/3} \cdot n^{-1/2}) = O(n^{1/6})$ going from plane binary trees to tanglegrams. Since the average height of a plane binary trees is asymptotically equal to $2\sqrt{\pi n}$ (see [14, Theorem MB]), the corollary follows.

□

Now, we consider the automorphism group. We call a generator of a binary tree $B$ a nonleaf vertex $v$ of $B$ such that the two subtrees stemming from $v$ are isomorphic and we denote by $sym(B)$ the number of such vertices. We remark that $|A(B)| = 2^{sym(B)}$, hence this also provides knowledge about the size of the automorphism group. The parameter $sym$ was studied (mean, limit law) by Bòna and Flajolet in [5] for plane binary trees so using Theorem 3.1.3, the results transfer to tanglegrams.

**Theorem 3.2.4** ([21])**.** *The expected number of generators of the top (bottom) half of a random tanglegram of size $n$ is asymptotically equal to $\lambda n$, where the constant $\lambda$, whose numerical value is* $0.2710416936\ldots$, *is the value of the function $f(x)$ defined by $f(x) = x + \frac{1}{2}f(x)^2 + (x - \frac{1}{2})f(x)^2$ at $x = \frac{1}{4}$. Moreover, the number of generators is asymptotically normally distributed.*

Again, the proof follows from the analogous statement for plane binary trees, see ([5, Theorem 2, (ii)]).
Lastly, as we have seen throughout the chapter, cherries play a major role, this fact was also mentioned in [3]. It is then natural to ask for the average number of matched cherries or more generally for the limiting distribution. This is stated in the next theorem.

**Theorem 3.2.5** ([21])**.** *The probability that there are exactly $k$ matched cherries in a random tanglegram of size $n$ converges to $e^{-1/4}4^{-k}/k!$, i.e. the number of matched cherries has a limiting Poisson distribution.*

*Proof.* Let $T$ be a tanglegram of size $n$. As in the proof of Corollary 3.2.2, the set of rotations of all $2(n-1)$ internal vertices of $T$ forms a group. This group acts on the set of all possible representations of this tanglegram as a pair of two plane binary trees with a bijection between the set of leaves. The number of orbits, which is then the number of distinct representations, is $2^{2(n-1)}/|A(T)|$ by the orbit-stabilizer theorem. Conversely, we can create a tanglegram $T$ from a pair of two plane binary trees $B_1, B_2$ with $n$ leaves and a bijection $\phi$ between the set of leaves. Therefore we assign a weight $|A(T)|/2^{2(n-1)}$ to the construction so that each distinct tanglegram is counted with a weight 1 when we sum over all possible choices of $B_1$, $B_2$ and $\phi$.
Now, we use this approach of counting tanglegrams as a way to estimate the probability of the event that there are exactly $k$ matched cherries. Let $B_1$ and $B_2$ be two plane binary trees and assume that they have $c_1$ and $c_2$ cherries, respectively. Afterwards, we count the number of bijections between the leaves of the two trees that create exactly $k$ matched cherries, where $k$ is a fixed nonnegative integer. Using the inclusion-exclusion principle, the number is given by

$$\binom{c_1}{k} \cdot \binom{c_2}{k} \cdot k!2^k \cdot \sum_{l \geq 0} (-1)^l \binom{c_1 - k}{l} \binom{c_2 - k}{l} \cdot l!2^l (n - 2k - 2l)!$$

$$= \frac{n!2^k}{k!} \sum_{l \geq 0} \frac{(-1)^l 2^l}{l!} \frac{\prod_{j=0}^{k+l-1} (c_1 - j)(c_2 - j)}{\prod_{j=0}^{2k+2l-1} (n - j)}.$$

On the left hand side of the equation, the quantities $\binom{c_1}{k}$ and $\binom{c_2}{k}$ account for the action of choosing $k$ cherries from each of the two trees. Then, $k!$ is the number of possible permutations of those $2k$ cherries so that we have $k$ matched cherries, and $2^k$ corrects the weight. Moreover, the sum takes care of the remaining cherries that are not matched (using inclusion-exclusion as mentioned earlier).

The sum is estimated in a similar way as in the proof of part (2) of Lemma 3.1.5. If we suppose that $c_1, c_2 \geq n/8$ (which we know to be the case for most choices of $B_1$ and $B_2$), then for $l \leq \sqrt{n}$, we have:

$$\frac{\prod_{j=0}^{k+l-1}(c_1 - j)(c_2 - j)}{\prod_{j=0}^{2k+2l-1}(n-j)} = \left(\frac{c_1 c_2}{n^2}\right)^{k+l}\Big(1 + O(l^2/n)\Big).$$

Furthermore, since $c_1, c_2 \leq n/2$, this fraction is bounded above by 1. Hence,

$$\begin{aligned}
&\sum_{l\geq 0} \frac{(-1)^l 2^l}{l!} \frac{\prod_{j=0}^{k+l-1}(c_1 - j)(c_2 - j)}{\prod_{j=0}^{2k+2l-1}(n-j)} \\
&= \sum_{0\leq l\leq\sqrt{n}} \frac{(-1)^l 2^l}{l!}\left(\frac{c_1 c_2}{n^2}\right)^{k+l}\Big(1 + O(l^2/n)\Big) + O\Big(\sum_{l>\sqrt{n}} \frac{2^l}{l!}\Big) \\
&= \sum_{l\geq 0} \frac{(-1)^l 2^l}{l!}\left(\frac{c_1 c_2}{n^2}\right)^{k+l} + O\Big(n^{-1}\sum_{l\geq 0}\frac{l^2 2^l}{l!}\left(\frac{c_1 c_2}{n^2}\right)^{k+l}\Big) + O\Big(\frac{2^{\lceil\sqrt{n}\rceil}}{\lceil\sqrt{n}\rceil!}\Big) \\
&= \left(\frac{c_1 c_2}{n^2}\right)^{k} \exp\Big(-\frac{2c_1c_2}{n^2}\Big) + O(n^{-1}).
\end{aligned}$$

If either $B_1$ or $B_2$ has fewer than $n/8$ cherries (which occurs with probability $O(n^{-1})$ if $B_1$ and $B_2$ are randomly selected), then we can bound the number of matchings with exactly $k$ matched cherries by $n!$. So as in the proof of Theorem 3.1.3, it follows that the number of triples of two plane binary trees and a bijection between their leaves such that there are exactly $k$ matched cherries is

$$C_{n-1}^2 \cdot n! \cdot \frac{2^k}{k!}\left(\mathbb{E}_n^{(P)}\Big(\Big(\frac{c(B_1)c(B_2)}{n^2}\Big)^k \exp\Big(-\frac{2c_1c_2}{n^2}\Big)\Big) + O(n^{-1})\right).$$

A Taylor expansion of $(x_1x_2)^k e^{-2x_1x_2}$ around $(\frac{1}{4}, \frac{1}{4})$ gives

$$16^{-k}e^{-1/8}\Big(1 + \frac{8k-1}{2}\Big(\Big(x_1 - \frac{1}{4}\Big) + \Big(x_2 - \frac{1}{4}\Big) + O\Big(\Big(x_1 - \frac{1}{4}\Big)^2 + \Big(x_2 - \frac{1}{4}\Big)^2\Big)\Big).$$

We will apply that to $x_1 = C(B_1)/n$ and $x_2 = C(B_2)/n$ as in the proof Theorem 3.1.3. Since under $\nu_n^{(P)}$, $c(B_1)$ and $c(B_2)$ are independent with mean $n(n-1)/(4n-6) = n/4 + O(1)$ and variance $O(n)$, we get:

$$\mathbb{E}_n^{(P)}\Big(\Big(\frac{c(B_1)c(B_2)}{n^2}\Big)^k \exp\Big(-\frac{2c_1c_2}{n^2}\Big)\Big) = 16^{-k}e^{-1/8} + O(n^{-1}).$$

Thus we find that the number of triples of two plane binary trees and a bijection between their leaves with exactly $k$ matched cherries is

$$C_n^2 \cdot n! \cdot \frac{e^{-1/8}}{8^k k!}(1 + O(n^{-1})).$$

Since permuting the leaves of those $k$ cherries constitutes $k$ rotations, the size of the automorphism group is at least $2^k$, so the associated weight is at least $2^k/2^{2(n-1)}$. Therefore, the probability that a random tanglegram of size $n$ has exactly $k$ matched cherries is at least

$$\frac{1}{t_n} \cdot C_n^2 \cdot n! \cdot \frac{e^{-1/8}}{8^k k!}(1 + O(n^{-1})) \cdot \frac{2^k}{2^{2(n-1)}} = \frac{e^{-1/4}4^{-k}}{k!}(1 + O(n^{-1})),$$

using the estimate for $t_n$ given in Corollary 2.3.2. Letting $p_{n,k}$ denote the probability that a random tanglegram of size $n$ has exactly $k$ matched cherries, we get:

$$\liminf_{n\to\infty} p_{n,k} \geq \frac{e^{-1/4}4^{-k}}{k!}.$$

The sum of these lower bounds is already 1, so $\lim_{n\to\infty} \sup p_{n,k}$ cannot be any greater. Therefore, we must have

$$\lim_{n\to\infty} p_{n,k} = \frac{e^{-1/4}4^{-k}}{k!},$$

which ends our proof. □

# Chapter 4

# Planar tanglegrams

It was pointed out in [7] that for applications, for instance in phylogenetics, we would like the tanglegram to have as few crossings as possible between inter-tree edges. This is called the *Tanglegram Layout* (TL) problem (see [22], [26]) or the *two-tree crossing minimization* problem referred to in [12]. This raises a natural question: among all tanglegrams, which of them can be drawn without crossings between inter-tree edges? More specifically, we established a formula for the number of tanglegrams of size $n$ in Chapter 2 so how many of them can be drawn without crossings? Here, we intend to enumerate those tanglegrams which we call *planar tanglegrams*. First, we will establish a surprising bijection between a special type of planar tanglegrams and pairs of triangulations without common diagonals. We will use this bijection to obtain several functional equations for the generating function of planar tanglegrams. Lastly, we will evaluate the number of planar tanglegrams using singularity analysis on their generating function. We note that this chapter is based on our joint work [28].

## 4.1 A bijection between planar tanglegrams and pairs of triangulations

**Definition 4.1.1.** A planar tanglegram is a tanglegram that can be drawn without crossings between inter-tree edges.

A planar tanglegram thus corresponds to a double coset of the form $A(T)IdA(S)$. Now, let $\mathcal{C}_b$ be the set of ordered pairs of rooted plane binary trees with the same number of leaves where we match one leaf of one tree to a leaf of the other without crossings. Using the crossing-free representation, every planar tanglegrams corresponds to an equivalence class of $\mathcal{C}_b$ (since several representations may be isomorphic).
We can go from one crossing-free representation of a planar tanglegram to another by successive rotations of internal vertices (the action of interchanging the two branches) using a similar argument as in the proof of Corollary 3.2.2 or Theorem 3.2.5 (see also [21, Theorem 8]). This will induce two types of planar tanglegrams: those that contain smaller tanglegrams

and those that do not. We make this more precise by defining a particular subgraph of a tanglegram.

**Definition 4.1.2.** A *binary subtree* $T'$ of a binary tree $T$ is an induced binary tree consisting of a vertex and all its successors. We call the binary subtree $T'$ a *proper binary subtree* if it is not a leaf and the root of $T'$ is different from the root of $T$.

**Definition 4.1.3.** A *subtanglegram* of a planar tanglegram consists of a binary subtree of the top tree and a binary subtree of the bottom tree with the same number of leaves, where each leaf of the top subtree is matched to a leaf of the bottom subtree. Moreover, a subtanglegram is called *proper subtanglegram* if the two corresponding binary subtrees are proper. Figure 4.1 shows a proper subtanglegram.

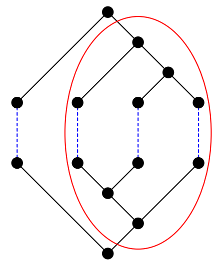

Figure 4.1: A proper subtanglegram.

Now, there are two types of planar tanglegrams: those that contain proper subtanglegrams and those that do not. This is also true for elements of $\mathcal{C}_b$ because these are, up to isomorphism, planar tanglegrams. Furthermore, there is a well-known bijection between plane binary trees and triangulations of polygons, so from a pair of plane binary trees, we can obtain a pair of triangulations. It will somehow be easier to work with pairs of triangulations than planar tanglegrams when we deal with the generating functions. The following theorem states that there is a bijection between elements of $\mathcal{C}_b$ and pairs of triangulations of a polygon and characterizes the property that an element of $\mathcal{C}_b$ contains a subtanglegram using a property of the pair of triangulations.

**Theorem 4.1.4.** *To every element $T$ of $\mathcal{C}_b$ with $n$ leaves corresponds a unique pair of triangulations of an $(n+1)$-gon. $T$ contains a proper subtanglegram if and only if the corresponding pair of triangulations has a common diagonal.*

*Proof.* We shall consider planted binary trees, we can obtain a rooted binary tree by deleting the pendent root. Let $T$ be a planted plane binary tree with $n$ leaves. We label the root 1 and the leaves by $\{2, \cdots, n+1\}$ as in Figure 4.2.

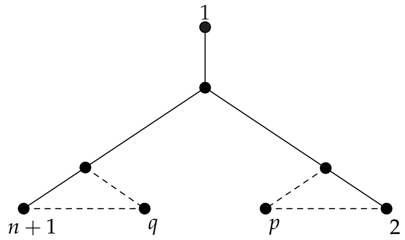


Figure 4.2: Tree numbering.

We extend the edges adjacent to the root and to any leaves to infinity in such a way that there is no intersection between those extended edges. Then, each vertex of degree 3 is surrounded by 3 regions in the plane as in Figure 4.3. We associate to each of these regions a unique point. For every pair of regions with a common border, we draw a curve between the corresponding points. At the end of the process, we obtain a triangulation of an $(n+1)$-gon where the sides and the diagonals might be curved (see Figure 4.4). This process is a bijection between planted binary trees and polygons since we are taking the geometric dual of a planar graph (see [17, p. 113-114]).

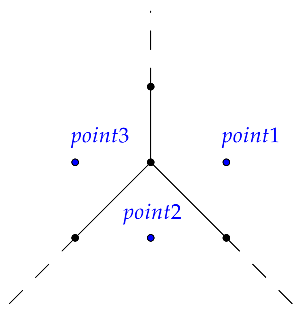


Figure 4.3: Regions.

Between any two leaves with consecutive numbers $j, j+1$ ($j \in \{1, \ldots, n\}$), there is a point corresponding to one region: we label this point $a_j$. Between the root and the leaf labelled $n+1$, there is another point: this point is labelled $a_{n+1}$. The points $a_j$ correspond to the vertices of the corresponding $(n+1)$-gon. The side of the polygon connecting $a_j$ and $a_{j+1}$ is labelled $j+1$ for $j \in \{1, \ldots, n\}$, and the side connecting $a_{n+1}$ and $a_1$ is labelled 1. This is illustrated by Figure 4.4 for $n = 5$.

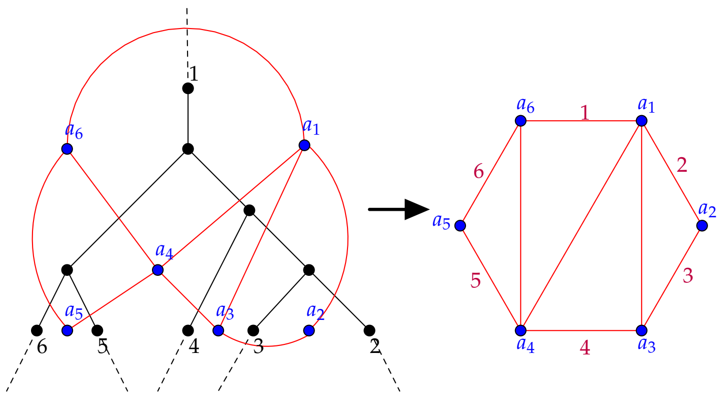


Figure 4.4: The correspondence between binary trees and triangulations.

Let $T_1$, $T_2$ be two planted plane binary trees with the same number of leaves $n$ numbered as in Figure 4.2. We flip the tree $T_2$ upside down and connect the leaves of $T_1$ to the leaves of $T_2$ with the same number. We obtain a pair of planted binary tree where the leaves are matched without crossings as in Figure 4.5. By deleting the two pendent roots, we obtain an element of $\mathcal{C}_b$, which is up to isomorphism a planar tanglegram; call this tanglegram $T_a$ (note that every element of $\mathcal{C}_b$ is obtained in this way).

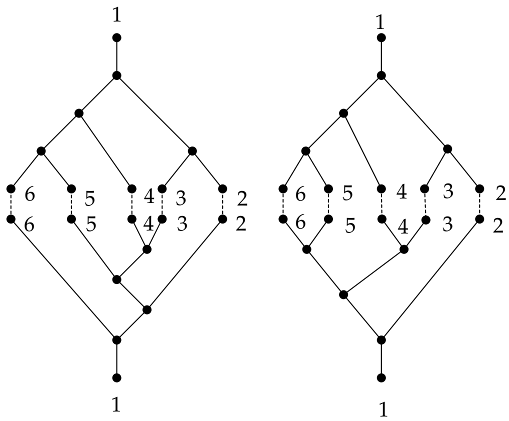


Figure 4.5: Examples of pairs of planted plane binary trees with matched leaves.

Now, suppose that $T_a$ contains a subtanglegram. Then there is a proper binary subtree $T_1'$ in $T_1$ and a proper binary subtree $T_2'$ in $T_2$ whose leaves are matched to $T_1'$. The proper binary subtrees $T_1'$ and $T_2'$ induce two sub-polygons of the polygons obtained respectively from $T_1$ and $T_2$.

Suppose that the labels of the leaves of $T_1'$ and $T_2'$ are from $p$ to $q$ ($p \leq q, p \geq 2, q \leq n+1$). Then, by construction, the corresponding vertices in the two triangulations from $a_{p-1}$ to $a_q$ are connected by edges (see Figure 4.6). Hence, the sides of the two sub-polygons from $a_{p-1}$ to $a_q$ have the same number.

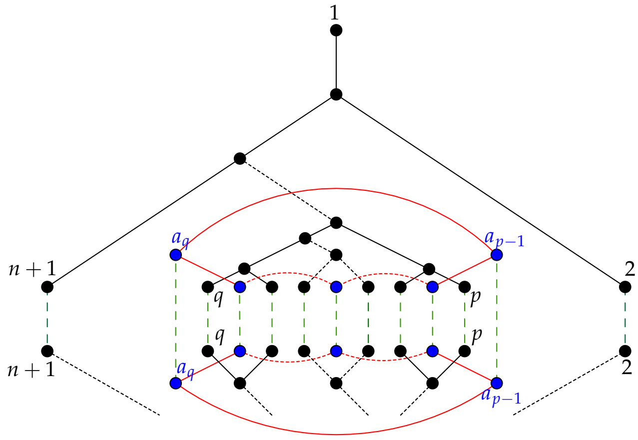


Figure 4.6: Subpolygons and subtanglegrams.

The edges connecting $a_q$ and $a_{p-1}$ in the polygons associated to $T_1'$ and $T_2'$ respectively are diagonals. Indeed, if $p = q$ then the subtanglegram is formed by only one leaf which is not a proper binary subtree. Also, in the case where $p = 2$ and $q = n+1$, we do not have a proper subtanglegram since the trees $T_1$ and $T_1'$ coincide, as do $T_2$ and $T_2'$. Thus the edge connecting $a_q$ and $a_{p-1}$ is a common diagonal of the two triangulations corresponding to the two trees $T_1$ and $T_2$.

Now, we consider the reverse process. The sides of the $C_{n-1}$ triangulations of an $(n+1)$-gon are labelled clockwise from 1 to $n+1$. Let $T_r$ be a triangulation of an $(n+1)$-gon. To each triangle in the triangulation, we associate one vertex. If two vertices are separated by a diagonal, then we connect them with an edge. If a side of a triangle is a side of the original polygon, then we draw an edge from the corresponding vertex of the triangle, cutting the edge of the polygon, to an external vertex outside of the polygon (see Figure 4.7).

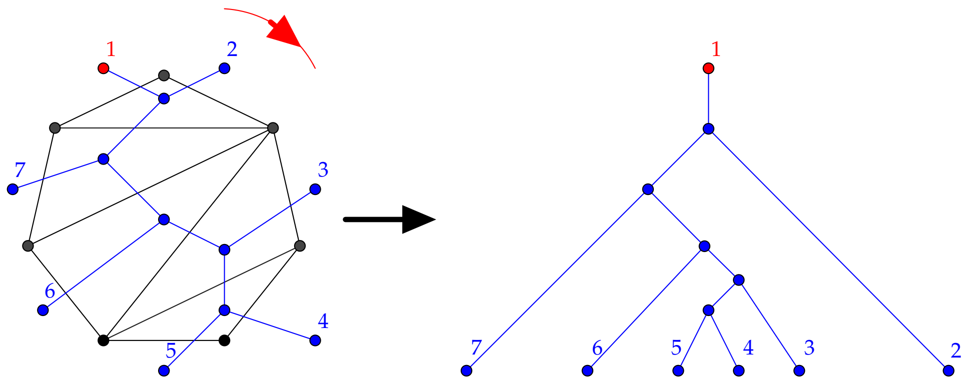


Figure 4.7: From a triangulation to a planted binary tree.

As a result of this process, we obtain a connected graph which has no cycle, i.e. a tree. Indeed, a cycle would have to enclose a vertex of the polygon, by construction, which is impossible. Note that in the reverse process, labeling the sides of the polygon is equivalent to labeling the

leaves of the tree. Hence the root corresponds to the side labeled 1 and we obtain a planted binary tree where the leaves are labeled from 2 to $n+1$ (see Figure 4.7).
Let $R_1$ and $R_2$ be two triangulations of an $(n+1)$-gon. We construct the tree corresponding to each triangulation $R_1$ and $R_2$. Then, we flip the second tree upside down, connect leaves with the same labels and obtain a tanglegram after deleting the roots (see Figure 4.8).

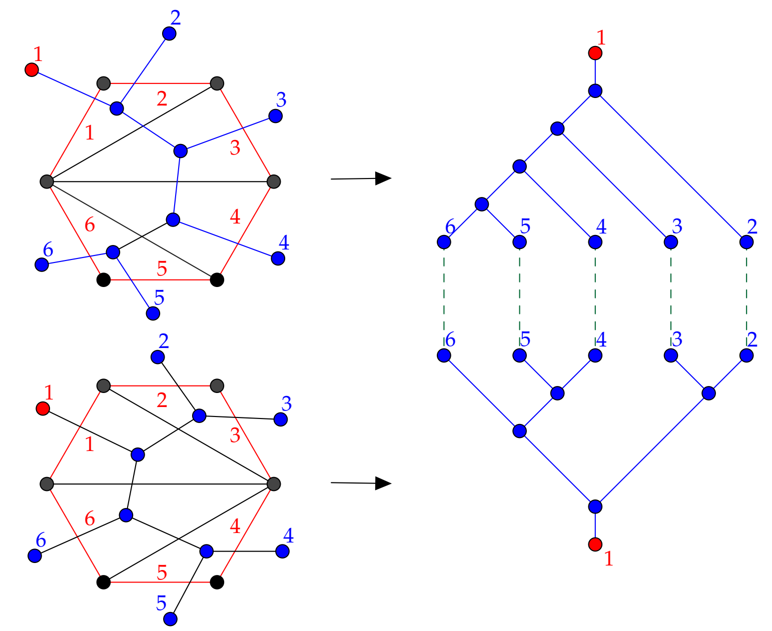

Figure 4.8: From a triangulation to a tanglegram.

Suppose that $R_1$ and $R_2$ have a common diagonal. This common diagonal divides each of the two polygons into two parts, one part containing the side that is labelled 1 while the other one does not. Consider the parts $R_1'$ and $R_2'$ of respectively $R_1$ and $R_2$ that do not contain 1. The sides of $R_1$ and $R_2$ have exactly the same labels. $R_1'$ and $R_2'$ are triangulations of a sub-polygon of $R_1$ and $R_2$ respectively. The common diagonal cuts an edge $e$ connecting an internal vertex of the part containing 1 and an internal vertex of $R_1'$ or $R_2'$ (see Figure 4.9).

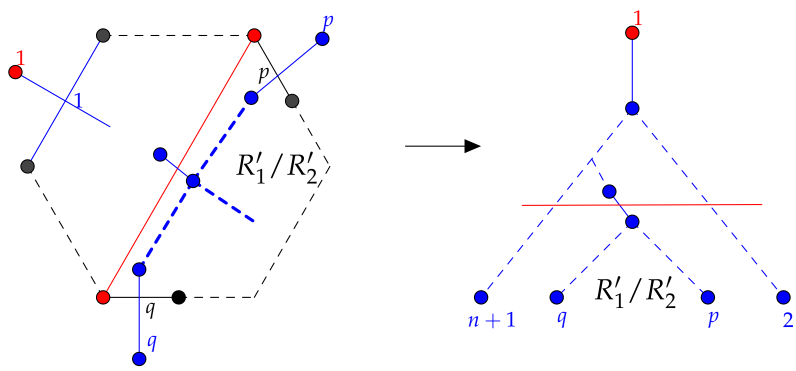


Figure 4.9: Sub-triangulations and binary subtrees.

By construction, considering only the sub-triangulations $R_1'$ and $R_2'$ is equivalent to cutting the edge $e$ and considering a proper binary subtree corresponding to $R_1'$ or $R_2'$ in the original tree (see Figure 4.9). Since the sides of $R_1'$ and $R_2'$ have the same labels, the leaves of the corresponding binary subtrees will have matching labels in the original trees. Thus, $R_1'$ and $R_2'$ form a proper subtanglegram of the original tanglegram. □

Now, we want to carry over the previous theorem to actual planar tanglegrams. We first give a definition for planar tanglegrams that do not contain proper subtanglegrams.

**Definition 4.1.5.** An *irreducible tanglegram* is a planar tanglegram which does not contain any proper subtanglegrams and which has more than one leaf. Furthermore, we call an element of $\mathcal{C}_b$ that corresponds to an irreducible tanglegram a *representation* of that tanglegram.

We extend Theorem 4.1.4 to irreducible tanglegrams.

**Theorem 4.1.6.** *We have the following properties:*

*(1) To every representation of an irreducible tanglegram with $n$ leaves corresponds a unique pair of triangulations of an $(n+1)$-gon without common diagonals.*

*(2) Every irreducible tanglegram with more than two leaves on each tree has precisely two possible representations, which are mirror images of each other.*

*(3) There is a bijection between irreducible tanglegrams and unordered pairs of triangulations that do not have a common diagonal.*

In order to prove Theorem 4.1.6, we first show that an irreducible tanglegram has a unique representation up to homeomorphism. This is done using a famous theorem of Whitney:

**Theorem 4.1.7** ([35]). *Every 3-connected planar graph has a unique plane embedding up to homeomorphism.*

We can apply Whitney's theorem to the graph obtained from a tanglegram by removing the leaves on each side (but leaving the connecting edges) and connecting the roots by an additional edge. Let us call this process *smoothing*. We have the following proposition:

**Proposition 4.1.8.** The graph obtained by smoothing a tanglegram is 3-regular and 3-connected if the tanglegram is irreducible and has more than 2 leaves on each tree.

*Proof.* First, notice that the process of smoothing an irreducible tanglegram does not create any parallel edges since the tanglegram would not be irreducible if that was the case (there would be a proper subtanglegram of size 2). After the process of smoothing, the remaining vertices (except the two roots) are all internal vertices, so they all have degree 3. The two roots are also of degree 3 because of the additional edge joining them. Thus, we have a 3-regular graph, which we will denote by $G$.

Next, let $T_1, T_2$ be the two halves of an irreducible planar tanglegram with more than two leaves on each side. We will prove that removing any pair of vertices $u, v$ of the graph obtained from the smoothing process does not disconnect the graph.

- Suppose $u, v$ are in the same tree, say $T_1$. Every vertex in $T_2$ is clearly still connected to $T_2$'s root. Every vertex in $T_1$ has three connections to the root of $T_2$ that are disjoint within $T_1$: via the root of $T_1$ and via the two children. Removing $u, v$ can only destroy at most two of them, so all vertices of $T_1$ are also still connected to the root of $T_2$. This means that $G - \{u, v\}$ is connected.

- Now, suppose that $u, v$ are in different trees. Assume that $u$ is a vertex of $T_1$, $v$ is a vertex of $T_2$ and that removing disconnects the graph obtained from the tanglegram by smoothing.

  $T_1 \setminus u$ has up to three components: two corresponding to the children of $u$, and one containing the root. Some of these components might be empty.

  Every non-empty component has at least two edges going to the other half of the tanglegram. Suppose there are only two, and both of them have $v$ as an end. Then we are in one of the following situations:

  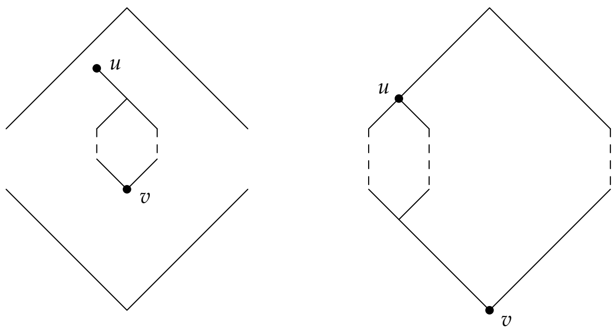
  

  Figure 4.10: Components of $G \setminus \{u, v\}$ with only 2 edges ending in $v$.

  Either way, there is a proper subtanglegram. So we can assume that every component of $T_1 \setminus u$ has an edge to $T_2 \setminus v$. The same applies to the components of $T_1 \setminus u$.

  Now consider the bipartite graph whose vertices are the components of $T_1 \setminus u$ and $T_2 \setminus v$, where we connect two components if there is an edge between them. If this graph is connected, then so is the graph $G \setminus \{u, v\}$. So call this graph $G'$ and suppose it is disconnected.

  Note that the root components of $T_1 \setminus u$ and $T_2 \setminus v$ (if they exist) are connected in $G'$ by definition (since there is an edge between the roots in $G$).

  So there must be a component of $G'$ containing only child components of $T_1 \setminus u$ and $T_2 \setminus v$ respectively. This component must have one of the following shapes, each corresponding to a proper subtanglegram:

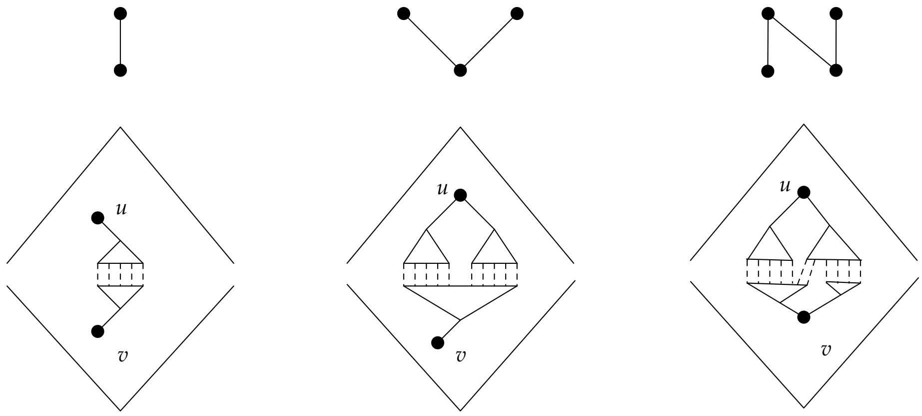


Figure 4.11: Component of $G'$ containing only child components of $T_1 \setminus u$ and $T_2 \setminus v$.

We conclude that $G$ is 3-connected.

□

*Proof of Theorem 4.1.6.* (1) The first part of Theorem 4.1.6 is a consequence of Theorem 4.1.4.

(2) Let $T_a$ be an irreducible tanglegram with more than two leaves in each tree, and let $T_1$ and $T_2$ be the corresponding binary trees. By Whitney's theorem and Proposition 4.1.8, there are only two possible representations, which are mirror images of each other. Suppose that the mirror image of $T_a$ is the same as $T_a$. Then, the mirror images of $T_1$ and $T_2$ are respectively the same as $T_1$ and $T_2$. This will imply that the left and right branches of $T_1$ and $T_2$ are the same. So the branches of $T_1$ and $T_2$ induce proper subtanglegrams since they contain more than one leaf each (see Figure 4.12). This contradict the fact that $T_a$ is irreducible. Thus, to an irreducible tanglegram with more than two leaves on each side correspond two distinct irreducible representations that are mirror images of each other.

(3) For $n = 2$ the statement is clearly true.

Now, suppose that $n > 2$. Let $P_n$ be the set of pairs of triangulations of an $(n+1)$-gon without common diagonal and let $I_n$ be the set of representations of irreducible tanglegrams with $n$ leaves. Moreover, denote by $P'_n$ the set of unordered pairs of triangulations of an $(n+1)$-gon without common diagonal and $I'_n$ the set of irreducible tanglegrams.

From the second part of the theorem, we know that to every element $T_a$ of $I'_n$, there are two distinct corresponding elements of $I_n$. So, to every pair of triangulations in $P'_n$, there are two distinct corresponding ordered pairs in $P_n$. This is because the two triangulations have to be distinct, as they would otherwise have a common diagonal.

By the first part of the theorem, there is a bijection $f$ from $P_n$ to $I_n$. Classifying the elements of $P_n$ by pairs that only differ by the order we obtain $P'_n$. Classifying the elements of $I_n$ by homeomorphic pairs (mirror image of each other) we get $I'_n$. Thus, the bijection $f$ induces a bijection $f'$ between $P'_n$ and $I'_n$.

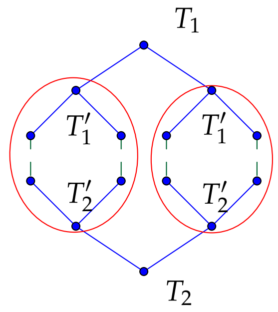


Figure 4.12: A tanglegram that is equal to its mirror image.

□

**Remark 4.1.9.** The only symmetric irreducible tanglegram (equal to its mirror image) is the tanglegram with two leaves in each tree.

## 4.2 Generating functions

Since there is a bijection between irreducible planar tanglegrams and unordered pairs of triangulations of a polygon without common diagonals, the generating functions of the two combinatorial objects are the same. Hence, we consider pairs of triangulations and we mark some of the common diagonals (not necessary all and possibly none) if there are any. This approach is described in [15, Section III.7]. We define a bivariate generating function $A(x, v)$ by:

$$A(x,v) = \sum_{(T_1,T_2)\in\mathcal{T}} \left( \sum_{m\in M(T_1,T_2)} v^{N(m)} \right) x^{n(T_1,T_2)}, \tag{4.2.1}$$

where $\mathcal{T}$ is the set of pairs of triangulations of polygons without marked diagonal. $M(T_1, T_2)$ is the set of all possible configurations of the pair $(T_1, T_2)$ that may contain marked diagonals; for $m \in M(T_1, T_2)$, we denote by $N(m)$ the number of marked diagonals in $m$. Lastly, $n(T_1, T_2)$ is the number of triangles in each of the two triangulations $T_1$ and $T_2$.
Let $T_1$ and $T_2$ be two triangulations of a polygon with $k(T_1, T_2)$ common diagonals. Then

$$\sum_{m\in M(T_1,T_2)} v^{N(m)} = (1+v)^{k(T_1,T_2)}.$$

Indeed, we can choose to mark a common diagonal, which yields a factor $v$ or not to mark it, which yields a factor 1. In addition, there are $k(T_1, T_2)$ diagonals to be marked. Thus,

$$A(x,v) = \sum_{(T_1,T_2)\in\mathcal{T}} \left( \sum_{m\in M(T_1,T_2)} v^{N(m)} \right) x^{n(T_1,T_2)} = \sum_{(T_1,T_2)\in\mathcal{T}} (1+v)^{k(T_1,T_2)} x^{n(T_1,T_2)}. \tag{4.2.2}$$

We define another generating function $B(x, v)$ by

$$B(x,v) = \sum_{(T_1,T_2)\in\mathcal{T}} \left( \sum_{m\in M'(T_1,T_2)} v^{N(m)} \right) x^{n(T_1,T_2)}, \tag{4.2.3}$$

where $M'(T_1, T_2)$ is the set of all possible configurations of $T_1$ and $T_2$ with no marked diagonals starting at the vertex labelled 1; $\mathcal{T}$, $N(m)$ and $n(T_1, T_2)$ are as in (4.2.1).

Let $\mathcal{A}$ be the set of all configurations consisting of two triangulations with some of the common diagonals potentially marked, and let $\mathcal{B}$ be the set of elements in $\mathcal{A}$ for which no marked diagonal starts at the vertex labelled one. Then $A(x, v)$, as defined in (4.2.2), is the bivariate generating function corresponding to $\mathcal{A}$ and $B(x, v)$ is the bivariate generating function corresponding to $\mathcal{B}$.

Now, let $A \in \mathcal{A}$. Suppose that $A$ contains $k$ marked diagonals starting at 1. Then $A$ can be decomposed into $k+1$ elements $B_1, B_2, \ldots, B_{k+1}$ in $\mathcal{B}$ separated by the marked diagonals. See Figure 4.13 for an illustration in the case $k = 4$.

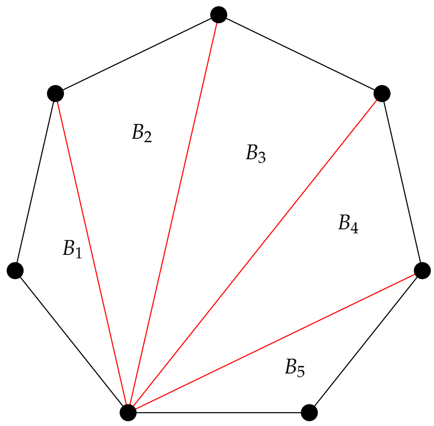


Figure 4.13: Triangulations with marked diagonals.

Using symbolic methods, this translates to:

$$\mathcal{A} \cong \bigcup_{k\geq 1} \{\bullet\}^{k-1} \times \mathcal{B}^k, \tag{4.2.4}$$

where $\bullet$ represents a marked diagonal. Thus we have

$$\begin{aligned} A(x,v) &= \sum_{k\geq 1} v^{k-1} B(x,v)^k \\ &= B(x,v) \sum_{k\geq 1} v^{k-1} B(x,v)^{k-1} \\ &= \frac{B(x,v)}{1 - vB(x,v)}. \end{aligned} \tag{4.2.5}$$

From the previous expression of $A(x, v)$ we get the following expression for $B(x, v)$:

$$B(x,v) = \frac{A(x,v)}{1 + vA(x,v)}. \tag{4.2.6}$$

Let $\mathcal{C}_r$ be the set of pairs of triangulations of polygons with $r$ triangles where we do not mark any diagonals. We can obtain an element $B$ of $\mathcal{B}$ by the following process: we take an element $C = (C_1, C_2)$ of $\mathcal{C}_r$ and several elements $A_1 = (A_1', A_1''), \ldots, A_k = (A_k', A_k'')$ of $\mathcal{A}$. We attach $A_1', A_2', \ldots, A_k'$ at different sides of the triangulation $C_1$ and $A_1'', A_2'', \ldots, A_k''$ at the corresponding sides of the triangulation $C_2$. These sides are chosen among those that are not adjacent to vertex 1. The sides where the $A_i'$ and $A_i''$ are attached becomes new marked diagonals. All other marked diagonals are taken from $A_1, A_2, \ldots, A_k$. Figure 4.14 shows the process for $C_1$ and $A_1'$. Since by construction $T$ does not contain any marked diagonals starting at 1, we have $T \in \mathcal{B}$.

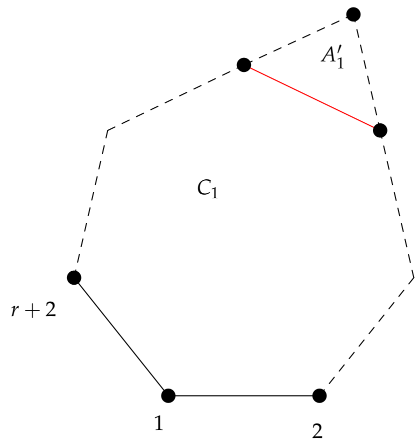


Figure 4.14: Triangulations without marked diagonal at 1.

Conversely, let $T \in \mathcal{B}$. Since there are no marked diagonals starting at vertex 1, there is a maximal sub-polygon that contain vertex 1 but no marked diagonal. The two triangulations of $T$ induce an element $C \in \mathcal{C}_r$ for some $r$. The remaining sub-polygons attached to the sides of $C$ will be elements of $\mathcal{A}$. By labelling those remaining elements of $\mathcal{A}$ from 1 to $k$ ($k \leq r$) we obtain the decomposition of $T$.
For a given $r$, each of the triangulations $C_1$, $C_2$ that form $C$ contains $r$ triangles, so both triangulations have $r+2$ sides. Since we are not allowed to attach elements of $\mathcal{A}$ to two sides adjacent to vertex 1, we have $r$ possible choices for the location of the $A_i$. Symbolically, the process translates to:

$$\mathcal{B} \cong \bigcup_{r \geq 1} \left( (\varnothing \cup (\mathcal{A} \times \{\bullet\}))^r \times \mathcal{C}_r \right). \tag{4.2.7}$$

Now, since we have $C_r$ triangulations of polygons with $r+2$ sides where $C_r = \frac{1}{r+1}\binom{2r}{r}$ is the $r^{\text{th}}$ Catalan number, translating (4.2.7) to generating functions gives us:

$$B(x, v) = \sum_{r \geq 1} (1 + A(x, v) \cdot v)^r \cdot x^r \cdot C_r^2. \tag{4.2.8}$$

From (4.2.6) and (4.2.8) we get

$$A(x,v) = \sum_{r\geq 1} x^r \cdot C_r^2 \cdot (1 + v \cdot A(x,v))^{r+1}. \tag{4.2.9}$$

Let $A(x) = A(x,-1)$ and $B(x) = B(x,-1)$. All the previous facts leads to the following theorem:

**Theorem 4.2.1.** *The generating function $H(x)$ of irreducible tanglegrams satisfies the following functional equations:*

$$H(x) = \frac{x \cdot A(x)}{2}, \tag{4.2.10}$$

$$A(x) = \sum_{r\geq 1} x^r \cdot C_r^2 \cdot (1 - A(x))^{r+1}. \tag{4.2.11}$$

**Remark 4.2.2.** The coefficient of $x^2$ in $H(x)$ is $\frac{1}{2}$. We maintain it as it is in order to take into account the symmetry when the tanglegram has two leaves.

*Proof.* From (4.2.2) we have:

$$A(x,v) = \sum_{(T_1,T_2)\in\mathcal{T}} (1+v)^{k(T_1,T_2)} x^{n(T_1,T_2)},$$

where $n(T_1,T_2)$ is the number of triangles and $k(T_1,T_2)$ is the number of common diagonals in $(T_1,T_2)$. By setting $v=-1$, all pairs of triangulations $(T_1,T_2)\in\mathcal{T}$ such that $k(T_1,T_2)\neq 0$ yield $(1+v)^{k(T_1,T_2)} = 0$. This means that all pairs of triangulations $(T_1,T_2)$ which have a common diagonal will not contribute to the sum for $A(x,-1)$. In another words, only the pairs of triangulations without common diagonal contribute to $A(x,-1) = A(x)$, i.e. $A(x)$ is the generating function of pairs of triangulations without common diagonals. Then, Equation (4.2.11) comes from Equation (4.2.9).

The coefficient of $x^r$ in $A(x)$ corresponds to pairs of triangulations without common diagonal with $r$ triangles in it. In addition, when we transform a triangulation of an $(r+2)$-gon into a planted binary tree, we obtain a planted binary tree with $r+1$ leaves. So, by part (1) of Theorem 4.1.6, the coefficient of $x^r$ in $A(x)$ is the number of representations of irreducible tanglegrams which have $r+1$ leaves. Multiplying $A(x)$ by $x$ gives us the generating function of representations of irreducible tanglegrams. From Theorem 4.1.6, we know that to every irreducible tanglegram with more than two leaves on each side, there are two irreducible representations. So, dividing $xA(x)$ by 2 gives us (4.2.10). □

The first few terms of $A(x)$ are given by

$$A(x) = x + 2x^2 + 10x^3 + 68x^4 + \cdots.$$

Now, the following theorem gives a functional equation relating the generating functions $T$ and $H$.

**Theorem 4.2.3.** *The generating function $T(x)$ of planar tanglegrams satisfies the following functional equation:*

$$T(x) = H(T(x)) + x + \frac{T(x^2)}{2}. \tag{4.2.12}$$

*Proof.* The term $x$ accounts for the tanglegram with only one leaf in each tree. Suppose we have a non-irreducible tanglegram $T_a$ with more than one leaf in each tree. It has proper maximal subtanglegrams (with respect to inclusion) $T_{a_1}, T_{a_2}, \ldots, T_{a_k}$ for some $k \in \mathbb{N}$ . For each subtanglegram $T_{a_i}$ we have two proper binary subtrees $T'_{a_i}$ and $T''_{a_i}$. We replace $T'_{a_i}$ and $T''_{a_i}$ by leaves and join the two leaves with an edge. By doing so, we obtain an irreducible tanglegram. Conversely, suppose that $T_a$ is an irreducible tanglegram with binary trees $T_1$ and $T_2$. Take each leaf $u \in T_1$ and the corresponding leaf $v \in T_2$ and replace $u, v$ together with the edge $\{u, v\}$ by a planar tanglegram with more than one leaf in each tree. Then, we obtain a non-irreducible tanglegram. This implies that we can obtain a planar tanglegram with more than one leaf on each side by replacing pairs of leaves of an irreducible tanglegram by any planar tanglegram (including the tanglegram with one leaf in each side). Then, we have two cases to consider:

- The irreducible tanglegram has more than two leaves on each side. Thus, the irreducible tanglegram is not symmetric. So, in order to obtain a new planar tanglegram, we replace an arbitrary pair of leaves with a planar tanglegram. Since in the monomial $x^r$ in $H(x)$, $r$ represents the number of leaves, replacing a pair of leaves by a planar tanglegram in the irreducible tanglegram translates to replacing $x$ by $T(x)$ in $H(x) - \frac{x^2}{2}$.

- The irreducible tanglegram has two leaves in each tree. We have to replace the two pairs of leaves in the irreducible tanglegram by two planar tanglegrams $T_{a_1}$, $T_{a_2}$ and take into account symmetries (unordered pairs of tanglegrams). By Pólya's enumeration theorem (see [18]) the generating function $C(x)$ of those unordered pairs is given by $C(x) = \frac{1}{2}\left(T(x)^2 + T(x^2)\right)$.

From the two previous cases, we get

$$\begin{aligned} T(x) &= x + \left(H(T(x)) - \frac{T(x)^2}{2}\right) + C(x) = x + \left(H(T(x)) - \frac{T(x)^2}{2}\right) + \frac{1}{2}\left(T(x)^2 + T(x^2)\right) \\ &= H(T(x)) + x + \frac{T(x^2)}{2}. \end{aligned}$$

□

The first 12 coefficients of $A(x)$, $H(x)$ and $T(x)$ are listed in the table below.

| $n$ | $[x^n]A(x)$ | $[x^n]H(x)$ | $[x^n]T(x)$ |
|---|---|---|---|
| 1 | 1 | 0 | 1 |
| 2 | 2 | 1/2 | 1 |
| 3 | 10 | 1 | 2 |
| 4 | 68 | 5 | 11 |
| 5 | 546 | 34 | 76 |
| 6 | 4872 | 273 | 649 |
| 7 | 46782 | 2436 | 6173 |
| 8 | 474180 | 23391 | 63429 |
| 9 | 5010456 | 237090 | 688898 |
| 10 | 54721224 | 2505228 | 7808246 |
| 11 | 613912182 | 27360612 | 91537482 |
| 12 | 7042779996 | 306956091 | 1102931565 |

Table 4.1: Table of the 12 first coefficients of $A(x), H(x)$ and $T(x)$.

## 4.3 Asymptotic analysis

In this section, we will estimate the number of irreducible tanglegrams using methods of analytic combinatorics. From relation (4.2.11), we have

$$\begin{aligned} A(x) &= \sum_{r\geq 1} x^r \cdot C_r^2 \cdot (1 - A(x))^{r+1} \\ &= (1 - A(x)) \sum_{r\geq 1} x^r \cdot C_r^2 \cdot (1 - A(x))^r. \end{aligned}$$

So,

$$A(x) = (1 - A(x)) \sum_{r\geq 1} C_r^2 \cdot (x(1 - A(x)))^r. \tag{4.3.1}$$

Let $u(x) = x(1 - A(x))$, so that $A(x) = 1 - \frac{u(x)}{x}$. Equation (4.3.1) becomes

$$1 - \frac{u(x)}{x} = \frac{u(x)}{x} \sum_{r\geq 1} C_r^2 \cdot u(x)^r. \tag{4.3.2}$$

Thus,

$$\begin{aligned} x &= u(x) + \sum_{r\geq 1} C_r^2 \cdot u(x)^{r+1} \\ &= \sum_{r\geq 0} C_r^2 \cdot u(x)^{r+1}. \end{aligned}$$

Consider the function $\phi(u) = \sum_{r\geq 0} C_r^2 \cdot u^{r+1}$. Then, $u(x)$ is the inverse function of $\phi$. Note that $\phi$ is a power series with radius of convergence $R = \frac{1}{16}$. In order to determine the analytic behaviour of $A(x)$, we have to investigate the analytic behaviour of the function

$$\phi(u) = \sum_{r\geq 0} C_r^2 \cdot u^{r+1} = \sum_{r=0}^{\infty} \frac{1}{(r+1)^2} \binom{2r}{r}^2 u^{r+1}.$$

We have the following theorem:

**Theorem 4.3.1.** *The function $\phi$ has an analytic continuation to the slit plane $\mathbb{C} \setminus [\frac{1}{16}, \infty)$. Moreover, when $u$ tends to $\frac{1}{16}$, we have*

$$\begin{aligned}\phi(u) &= \frac{4-\pi}{4\pi} - \frac{1}{4\pi}(1-16u)\\ &- \frac{1}{64\pi}\Big(5 - 8\log 2 + 2\log(1-16u)\Big)(1-16u)^2 + O\Big(\Big|(1-16u)^3 \log(1-16u)\Big|\Big).\end{aligned}$$

*Proof.* First of all, it is well known that

$$\sum_{r=0}^{\infty} \binom{2r}{r}^2 u^r = \frac{2}{\pi} k(16u) \tag{4.3.3}$$

for $|u| < \frac{1}{16}$, with the complete elliptic integral

$$k(x) = \int_0^{\pi/2} \frac{1}{\sqrt{1 - x\sin^2 t}}\, dt. \tag{4.3.4}$$

The integral $k(x)$ defines an analytic function on the slit plane $\mathbb{C} \setminus [1, \infty)$. Now, [9, 19.12.1] (online companion of [24]) gives the series

$$k(1-x) = \sum_{m=0}^{\infty} \binom{m-1/2}{m}^2 x^m \Big( -\frac{1}{2}\log x + d(m) \Big)$$

after some rewriting (note that [9] uses a different notation, where $K(u) = k(u^2)$ according to our notation), where

$$d(0) = 2\log 2, \quad d(m) = d(m-1) - \frac{1}{m(2m-1)},$$

or equivalently $d(m) = \psi(1+m) - \psi(\frac{1}{2} + m)$ (here $\psi$ is the Digamma function). Now since

$$\sum_{m=0}^{\infty} \binom{m-1/2}{m}^2 x^m = \frac{2}{\pi} k(x),$$

which is in fact equivalent to (4.3.3), we can also write this as

$$k(1-x) = -\frac{1}{\pi} k(x) \log x + \sum_{m=0}^{\infty} \binom{m-1/2}{m}^2 d(m) x^m,$$

which provides us with an analytic continuation around the branch cut for $|x| < 1$, $x \notin (-1, 0]$. In particular, we have the following asymptotic expansion around $u = \frac{1}{16}$ (by taking the first term in the series):

$$k(16u) = 2\log 2 - \frac{1}{2}\log(1-16u) + O\Big(\Big|(1-16u)\log(1-16u)\Big|\Big).$$

Now

$$\phi(u) = \frac{2}{\pi}\int_0^u \frac{1}{v}\int_0^v k(16z)\,dz\,dv,$$

which provides an analytic continuation of $\phi$ to the slit plane $\mathbb{C} \setminus [\frac{1}{16}, \infty)$. The asymptotic expansion can be integrated termwise by writing

$$\int_0^v k(16z)\,dz = \int_0^{1/16} k(16z)\,dz - \int_v^{1/16} k(16z)\,dz,$$

cf. [15, TheoremVI.9]. All we need is first the value

$$\int_0^{1/16} k(16z)\,dz = \frac{1}{8}$$

and later

$$\phi\Big(\frac{1}{16}\Big) = \frac{2}{\pi}\int_0^{1/16}\frac{1}{v}\int_0^v k(16z)\,dz\,dv = \frac{4-\pi}{4\pi},$$

which can be obtained by plugging in (4.3.4) and interchanging the order of integration. This gives us first

$$\int_0^v k(16z)\,dz = \frac{1}{8} - \frac{1}{32}\Big(1 + 4\log 2 - \log(1-16v)\Big)(1-16v) + O\Big(\Big|(1-16v)^2\log(1-16v)\Big|\Big).$$

Then, by multiplication with

$$\frac{1}{v} = 16 + 16(1-16v) + O\big(|1-16v|^2\big),$$

we obtain

$$\frac{1}{v}\int_0^v k(16z)\,dz = 2 + \frac{1}{2}\Big(3 - 4\log 2 + \log(1-16v)\Big)(1-16v) + O\Big(\Big|(1-16v)^2\log(1-16v)\Big|\Big).$$

Since we have

$$\begin{aligned}\phi(u) &= \frac{2}{\pi}\int_0^{1/16}\frac{1}{v}\int_0^v k(16z)\;dz\;dv + \frac{2}{\pi}\int_{1/16}^u \frac{1}{v}\int_0^v k(16z)\;dz\;dv\\ &= \phi\Big(\frac{1}{16}\Big) + \frac{2}{\pi}\int_{1/16}^u \frac{1}{v}\int_0^v k(16z)\;dz\;dv.\end{aligned}$$

One more integration step yields

$$\begin{aligned}\phi(u) = \;&\frac{4-\pi}{4\pi} - \frac{1}{4\pi}(1-16u)\\ &- \frac{1}{64\pi}\Big(5 - 8\log 2 + 2\log(1-16u)\Big)(1-16u)^2 + O\Big(\Big|(1-16u)^3\log(1-16u)\Big|\Big).\end{aligned}$$

$\square$

In order to determine the asymptotic behavior of $A$ (thus the asymptotic behavior of $H$), we will need the asymptotic behavior of $u$. The latter implies that we need to be able to invert $\phi$ in an appropriate region of $\mathbb{C}$. The next lemma will help us with that.

**Lemma 4.3.2.** $\phi$ is injective in $\mathbb{C} \setminus [\frac{1}{16}, \infty)$ and for all $u \in \mathbb{C} \setminus [\frac{1}{16}, \infty)$ we have $\phi'(u) \neq 0$.

*Proof.* We notice that:

$$\phi'(u) = \frac{1}{4\pi u} \int_0^{\frac{\pi}{2}} \frac{1 - \sqrt{1 - 16u\sin^2(t)}}{\sin^2(t)}\, dt = \frac{4}{\pi} \int_0^{\frac{\pi}{2}} \frac{1}{1 + \sqrt{1 - 16u\sin^2(t)}}\, dt,$$

which follows from differentiating (4.3.4). Now we can make use of the fact that

$$Re\left(1 + \sqrt{1 - 16u\sin^2(t)}\right) > 0,$$

since $Re(\sqrt{z}) > 0$ holds for every $z \in \mathbb{C} \setminus (-\infty, 0]$. Thus,

$$Re\left(\frac{1}{1 + \sqrt{1 - 16u\sin^2(t)}}\right) > 0$$

which in turn means that $Re(\phi'(u)) > 0$ for all $u \in \mathbb{C} \setminus [\frac{1}{16}, \infty)$. In particular, $\phi'(u) \neq 0$ for all possible values of $u$. In the same way, we can show that $Im(\phi'(u))$ has the same sign as $Im(u)$ for all $u$, and the two combined imply that $\phi$ is injective on its domain of analyticity. Indeed, let $u, v \in \mathbb{C}$ such that $u \neq v$. Since $\phi$ is analytic in $\mathbb{C} \setminus [\frac{1}{16}, \infty)$, we have

$$\phi(u) = \phi(v) + \int_u^v \phi'(z) dz,$$

where we can integrate along any path joining $u$ and $v$ in the slit plane. We have several cases to consider:

- Suppose $u, v$ are in different half-planes.

  *Claim:* $Im(\phi(u))$ has the same sign as $Im(u)$.

  It is enough to show this for $Im(u) > 0$ since we have $\phi(z) = \overline{(\phi(\bar{z}))}$. So let $u = u_1 + iu_2$, $u_2 > 0$ and $v = v_1 + iv_2$. We have two cases to consider:

  – If $u_1 \geq 0$ then

$$\begin{aligned}
\phi(u) &= \int_0^u \phi'(z)\, dz \\
&= \int_0^{iu_2} \phi'(z)\, dz + \int_{iu_2}^{u_1 + iu_2} \phi'(z)\, dz \\
&= i\int_0^{u_2} \phi'(iw)\, dw + \int_0^{u_1} \phi'(t + iu_2)\, dt.
\end{aligned}$$

  Thus,

$$Im(\phi(u)) = \int_0^{u_2} Re(\phi'(iw))\, dw + \int_0^{u_1} Im(\phi'(t + iu_2))\, dt > 0.$$

- If $u_1 < 0$ then

$$\phi(u) = \phi(u_1) + \int_{u_1}^{u_1+iu_2} \phi'(z)dz = \phi(u_1) + i\int_0^{u_2} \phi'(u_1+it)dt.$$

Hence $Im(\phi(u)) = \int_0^{u_2} Re(\phi'(u_1+it))dt > 0$, which proves the claim.

Therefore, $\phi(u) \neq \phi(v)$ if $Im(u)$ and $Im(v)$ have different signs.

- Suppose that $u$ and $v$ are in the same half plane. Let $u = u_1 + iu_2$ and $v = v_1 + iv_2$ such that $u_1 \leq v_1$ and $u_2, v_2 \geq 0$ without loss of generality since $\phi(\bar{z}) = \overline{\phi(z)}$. Then, we have:

$$\begin{aligned}\phi(v) - \phi(u) &= \int_u^v \phi'(z)\, dz \\ &= \int_{u_1+iu_2}^{u_1+iv_2} \phi'(z)\, dz + \int_{u_1+iv_2}^{v_1+iv_2} \phi'(z)dz \\ &= i\int_{u_2}^{v_2} \phi'(u_1+iw)\, dw + \int_{u_1}^{v_1} \phi'(t+iv_2)\, dt.\end{aligned}$$

  - If $u_2 \leq v_2$, we consider the imaginary part:

$$Im(\phi(v) - \phi(u)) = \int_{u_2}^{v_2} Re(\phi'(u_1+iw))\, dw + \int_{u_1}^{v_1} Im(\phi'(t+iv_2))\, dt > 0.$$

  - If $u_2 \geq v_2$, we consider the real part:

$$Re(\phi(v) - \phi(u)) = -\int_{u_2}^{v_2} Im(\phi'(u_1+iw))\, dw + \int_{u_1}^{v_1} Re(\phi'(t+iv_2))\, dt > 0.$$

  In either case, $\phi(u) \neq \phi(v)$.

We conclude that $\phi(u) \neq \phi(v)$ in all cases and that $\phi$ is injective. □

Now, we have the following asymptotic formula for $H$ which will give us an asymptotic formula for the number of irreducible tanglegrams.

**Theorem 4.3.3.** *There exist constants $\theta \in (0, \frac{\pi}{2})$ and $\epsilon > 0$ such that $H$ is analytic in*

$$\Delta = \{x \mid |x| < \alpha + \epsilon \text{ and } |Arg(x - \alpha)| > \theta\},$$

*and for $x \in \Delta$, we have*

$$H(x) = C_0 + C_1(\alpha - x) + C_2(\alpha - x)^2 + B(\alpha - x)^2 \log(\alpha - x) + O(|(\alpha - x)^3 \log(\alpha - x)|) \quad (4.3.5)$$

*where*

- $C_0 = \frac{16-5\pi}{32\pi} = \frac{\alpha}{2} - \frac{1}{32} = H(\alpha)$,

- $C_1 = \frac{\pi}{8} - \frac{1}{2} = -H'(\alpha)$,
- $C_2 = -\frac{\pi^2}{32}(5 - 8\log 2 + 2\log 4\pi)$,
- $B = -\frac{\pi^2}{16}$ *and*
- $\alpha = \phi(1/16) = \frac{4-\pi}{4\pi}$.

*Thus, the number of irreducible planar tanglegrams is asymptotically:*

$$[x^n]H(x) \sim \frac{(\pi\alpha)^2}{8} \cdot n^{-3} \cdot \alpha^{-n}.$$

*Here* $\alpha \approx 0.0683$ *and* $\alpha^{-1} \approx 14.6391$.

*Proof.* In order to simplify computations, we write $y = 1 - 16u$. We let $u$ tend to $\frac{1}{16}$ so $y$ tends to 0. Then, by Theorem 4.3.1, we have

$$\phi(u) = x = \alpha - \frac{1}{4\pi}y - \frac{1}{32\pi}y^2\log(y) + O(|y^2|). \tag{4.3.6}$$

By Lemma 4.3.2 and the implicit function theorem, $\phi$ is invertible and the inverse $\phi^{-1}$ is analytic in $\phi\Big(\mathbb{C} \setminus [\frac{1}{16}, \infty)\Big)$. The function $\phi$ comes from the integration of $k$, which has a branch cut of square root type. The cut $[\frac{1}{16}, \infty)$ is mapped to two branches (see Figure 4.15). In view of the expansion (4.3.6), it is possible to choose $\theta$ in such a way that $\Delta$ lies in the image $\phi(\mathbb{C} \setminus [\frac{1}{16}, \infty))$. Hence, $u = \phi^{-1}$ is well defined and analytic in $\Delta$.

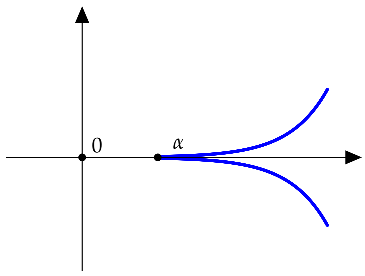


Figure 4.15: Two branches of $\phi$.

Then, from Equation (4.3.6), we have

$$4\pi(\alpha - x) = y + \frac{1}{8}y^2\log(y) + O(|y|^2).$$

By bootstrapping, we will establish an asymptotic expansion of $y$ with an error of

$$O\Big(\Big|(\alpha - x)^3\log(\alpha - x)\Big|\Big).$$

First, we will determine an asymptotic expansion with an error of $O(|(\alpha - x)^2|)$. If $y$ tends to 0 then $x$ tends to $\alpha$ and $\frac{4\pi(\alpha - x)}{y}$ tends to 1. Thus $y \sim 4\pi(\alpha - x)$ i.e.

$$y = 4\pi(\alpha - x) + y_1(x),$$

where $y_1(x) = O\Big(\Big|(\alpha - x)^2 \log(\alpha - x)\Big|\Big)$.
Plugging the previous expression for $y$ into Equation (4.3.6) gives us

$$0 = -\frac{1}{4\pi}y_1(x) - \frac{1}{32\pi}\Big(4\pi(\alpha - x) + y_1(x)\Big)^2 \log\Big(4\pi(\alpha - x) + y_1(x)\Big) + O\Big(\Big|\Big(4\pi(\alpha - x) + y_1(x)\Big)^2\Big|\Big). \tag{4.3.7}$$

Since $y_1(x) = O\Big(\Big|(\alpha - x)^2 \log(\alpha - x)\Big|\Big)$ then

$$y_1(x) \cdot 4\pi(\alpha - x) = O\Big(\Big|(\alpha - x)^3 \log(\alpha - x)\Big|\Big)$$

and

$$y_1^2(x) = O\Big(\Big|(\alpha - x)^4 \log^2(\alpha - x)\Big|\Big).$$

The main term in $\Big(4\pi(\alpha - x) + y_1(x)\Big)^2$ is $16\pi^2(\alpha - x)^2$. Next, we have

$$\log\Big(4\pi(\alpha - x) + y_1(x)\Big) = \log(4\pi) + \log(\alpha - x) + \log\Big(1 + \frac{y_1(x)}{4\pi(\alpha - x)}\Big).$$

Therefore,

$$(\alpha - x)^2 \log\Big(4\pi(\alpha - x) + y_1(x)\Big) = C(\alpha - x)^2 + (\alpha - x)^2 \log(\alpha - x) + O\Big(\Big|(\alpha - x)^3 \log(\alpha - x)\Big|\Big)$$

where $C = \log 4\pi$. So from Equation (4.3.7) we have:

$$y_1(x) = -2\pi^2(\alpha - x)^2 \log(\alpha - x) + O\Big(\Big|(\alpha - x)^2\Big|\Big).$$

Hence,

$$y = 4\pi(\alpha - x) - 2\pi^2(\alpha - x)^2 \log(\alpha - x) + O\Big(\Big|(\alpha - x)^2\Big|\Big). \tag{4.3.8}$$

Now, we will determine an asymptotic expansion of order $O\Big(\Big|(\alpha - x)^3 \log(\alpha - x)\Big|\Big)$. Again, from Theorem 4.3.1, we have:

$$x = \alpha - \frac{1}{4\pi}y - \frac{1}{64\pi}(K + 2\log(y))y^2 + O\Big(\Big|(\alpha - x)^3 \log(\alpha - x)\Big|\Big). \tag{4.3.9}$$

where $K = 5 - 8\log 2$. Then, from Equation (4.3.8), we have

$$y = 4\pi(\alpha - x) - 2\pi^2(\alpha - x)^2 \log(\alpha - x) + y_2(x), \tag{4.3.10}$$

where $y_2(x) = O(|(x - \alpha)^2|)$. Then, plugging Equation (4.3.10) into Equation (4.3.9), we get

$$\begin{aligned} x = \alpha &- \frac{1}{4\pi}\Big(4\pi(\alpha - x) - 2\pi^2(\alpha - x)^2 \log(\alpha - x) + y_2(x)\Big) \\ &- \frac{1}{64\pi}\Big(K + 2\log\Big(4\pi(\alpha - x) - 2\pi^2(\alpha - x)^2 \log(\alpha - x) + y_2(x)\Big)\Big) \\ &\Big(4\pi(\alpha - x) - 2\pi^2(\alpha - x)^2 \log(\alpha - x) + y_2(x)\Big)^2. \end{aligned} \tag{4.3.11}$$

Let

$$M(x) = \log\Big(4\pi(\alpha-x) - 2\pi^2(\alpha-x)^2\log(\alpha-x) + y_2(x)\Big)\Big)$$

and

$$N(x) = \Big(4\pi(\alpha-x) - 2\pi^2(\alpha-x)^2\log(\alpha-x) + y_2(x)\Big)^2.$$

Then

$$\begin{aligned}M(x) &= \log(4\pi) + \log(\alpha-x) + \log\Big(1 - \frac{\pi}{2}(\alpha-x)\log(\alpha-x) + \frac{y_2(x)}{4\pi(\alpha-x)}\Big)\\ &= \log(4\pi) + \log(\alpha-x) + \frac{\pi}{2}(\alpha-x)\log(\alpha-x) + O(|(\alpha-x)^2|)\end{aligned}$$

and

$$\begin{aligned}N(x) &= 16\pi^2(\alpha-x)^2 - 16\pi^3(\alpha-x)^3\log(\alpha-x) + 8\pi(\alpha-x)y_2(x)\\ &\quad + 4\pi^4(\alpha-x)^4\log^2(\alpha-x) - 4\pi^2 y_2(x)(\alpha-x)^2\log(\alpha-x) + y_2^2(x).\end{aligned}$$

Since $y_2(x) = O(|(\alpha-x)|^2)$, we have $N(x) = 16\pi^2(\alpha-x)^2 + O\Big(\Big|(\alpha-x)^3\log(\alpha-x)\Big|\Big)$. Hence, from Equation (4.3.11), we get

$$\begin{aligned}0 &= \frac{\pi}{2}(\alpha-x)^2\log(\alpha-x) - \frac{1}{4\pi}y_2(x) - \frac{K}{64\pi}\Big(16\pi^2(\alpha-x)^2 + O\Big(\Big|(\alpha-x)^3\log(\alpha-x)\Big|\Big)\Big)\\ &- \frac{1}{32\pi}\Big(\log(4\pi) + \log(\alpha-x) + \frac{\pi}{2}(\alpha-x)\log(\alpha-x) + O(|(\alpha-x)^2|)\Big)\\ &\Big(16\pi^2(\alpha-x)^2 + O\Big(\Big|(\alpha-x)^3\log(\alpha-x)\Big|\Big)\Big).\end{aligned}$$

Thus,

$$\begin{aligned}0 &= \frac{\pi}{2}(\alpha-x)^2\log(\alpha-x) - \frac{1}{4\pi}y_2(x)\\ &- \Big(\frac{\pi}{4}K + \frac{\pi}{2}\log(4\pi)\Big)(\alpha-x)^2 - \frac{\pi}{2}(\alpha-x)^2\log(\alpha-x) + O\Big(\Big|(\alpha-x)^3\log(\alpha-x)\Big|\Big).\end{aligned}$$

So,

$$y_2(x) = -\pi^2(K + 2\log(4\pi))(\alpha-x)^2 + O\Big(\Big|(\alpha-x)^3\log(\alpha-x)\Big|\Big)$$

and therefore

$$y = 4\pi(\alpha-x) - 2\pi^2(\alpha-x)^2\log(\alpha-x) - \pi^2(K+2\log(4\pi))(\alpha-x)^2 + O\Big(\Big|(\alpha-x)^3\log(\alpha-x)\Big|\Big).$$

Since $y = 1 - 16u$, we have

$$\begin{aligned}u &= \frac{1}{16} - \frac{\pi}{4}(\alpha-x) + \frac{\pi^2}{8}(\alpha-x)^2\log(\alpha-x)\\ &+ \frac{\pi^2}{16}(K+2\log(4\pi))(\alpha-x)^2 + O\Big(\Big|(\alpha-x)^3\log(\alpha-x)\Big|\Big).\end{aligned}$$

Finally, $u = x(1 - A(x))$ so

$$
\begin{aligned}
H(x) = \frac{xA(x)}{2} &= \frac{x}{2} - \frac{1}{32} + \frac{\pi}{8}(\alpha - x) - \frac{\pi^2}{16}(\alpha - x)^2 \log(\alpha - x) - \frac{\pi^2}{32}(K + 2\log(4\pi))(\alpha - x)^2 \\
&+ O\Big(\Big|(\alpha - x)^3 \log(\alpha - x)\Big|\Big) \\
&= H(\alpha) - H'(\alpha)(\alpha - x) + C_2(\alpha - x)^2 + B(\alpha - x)^2 \log(\alpha - x) + O\Big(\Big|(\alpha - x)^3 \log(\alpha - x)\Big|\Big) \\
&= C_0 + C_1(\alpha - x) + C_2(\alpha - x)^2 + B(\alpha - x)^2 \log(\alpha - x) + O\Big(\Big|(\alpha - x)^3 \log(\alpha - x)\Big|\Big).
\end{aligned}
$$

We have

$$
-(\alpha - x)^2 \log(\alpha - x) = \alpha^2 \Big(1 - \frac{x}{\alpha}\Big)^2 \log\Big(\frac{1}{1 - \frac{x}{\alpha}}\Big) - \alpha^2 \log(\alpha) \Big(1 - \frac{x}{\alpha}\Big)^2.
$$

So,

$$
H(x) = C_0 + C_1(\alpha - x) - B\alpha^2 \Big(1 - \frac{x}{\alpha}\Big)^2 \log\Big(\frac{1}{1 - \frac{x}{\alpha}}\Big) + K'(1 - \frac{x}{\alpha})^2 + O\Big(\Big|\Big(1 - \frac{x}{\alpha}\Big)^3 \log\Big(1 - \frac{x}{\alpha}\Big)\Big|\Big)
$$

where $K' = \alpha^2(C_2 + B\log(\alpha))$.

When $n > 2$, the coefficient of $x^n$ in $(1 - \frac{x}{\alpha})^2$ vanishes, so asymptotically only $\left(1 - \frac{x}{\alpha}\right)^2 \log\left(\frac{1}{1-\frac{x}{\alpha}}\right)$ contributes to $[x^n]H(x)$ and the error is of order $O\Big(\Big|\left(1 - \frac{x}{\alpha}\right)^3 \log\left(1 - \frac{x}{\alpha}\right)\Big|\Big)$.

In order to determine the asymptotic behaviour of the coefficients of $\left(1 - \frac{x}{\alpha}\right)^2 \log\left(\frac{1}{1-\frac{x}{\alpha}}\right)$, we consider the function $f(z) = (1 - z)^{-2} \log\left(\frac{1}{1-z}\right)$. By [15, TheoremVI.2, Special cases], we have

$$
[z^n]f(z) = \frac{2}{n(n-1)(n-2)} \sim 2n^{-3}.
$$

The function $H$ and the region $\Delta$ satisfy the conditions of [15, TheoremVI.4], so we have:

$$
[x^n]H(x) \sim -2 \cdot \alpha^2 \cdot B \cdot \alpha^{-n} \cdot n^{-3} = \frac{(\pi\alpha)^2}{8} \cdot n^{-3} \cdot \alpha^{-n}.
$$

□

The last step is to study the asymptotic behavior of the generating function of planar tanglegrams.

**Remark 4.3.4.** From Equation (4.2.12) we have

$$
T(x) = H(T(x)) + x + \frac{T(x^2)}{2}.
$$

We investigate each term on the right side of the previous equation. First, $x$ is an entire function. Next, if we let $\rho$ be the radius of convergence of $T(x)$, then $\rho \leq \alpha < 1$ since the coefficients of $T$ are greater than or equal to the coefficients of $H$. Since the radius of convergence of $T(x^2)$ is $\sqrt{\rho}$ it follows that $T(x^2)$ has a radius of convergence greater than $T(x)$. Thus, the dominant singularity of $T(x)$ is inherited from the dominant singularity of $H(x)$.

The previous remark leads to the following proposition:

**Proposition 4.3.5.** $T(x)$ is analytic in an open disk centered at 0 with radius $\rho > 0$ and the dominant singularity $\rho$ of $T(x)$ satisfies $T(\rho) = \alpha = \frac{4-\pi}{4\pi}$.

*Proof.* $T(x)$ has non-negative coefficients, so by Pringsheim's Theorem, the radius of convergence $\rho$ of $T(x)$ is also a singularity. We know that $H(x)$ has its dominant singularity at $x = \alpha = \frac{4-\pi}{4\pi}$, so $H(T(x))$ has a singularity at any point $x$ for which $T(x) = \alpha = \frac{4-\pi}{4\pi}$. Suppose that there exists a positive real number $\tau$ such that $\tau < \alpha$ and $H(T(x))$ is singular at $T(x_0) = \tau$ for some $x_0 > 0$. We define the bivariate function

$$F(t,x) = H(t) + x + \frac{T(x^2)}{2} - t,$$

and we have

$$\frac{\partial F(t,x)}{\partial t} = H'(t) - 1.$$

Since $\tau$ is a singularity of $T(x)$, the implicit function theorem has to fail at $(t,x) = (\tau, x_0)$ for $F(t,x)$. In other words, we must have

$$\frac{\partial F(t,x)}{\partial t}(\tau, x_0) = H'(\tau) - 1 = 0 \text{ i.e. } H'(\tau) = 1.$$

Next, we have

$$H'(x) = \frac{A(x) + xA'(x)}{2}.$$

So, $H'(x)$ has non-negative coefficients. Hence $H'(x)$ is an increasing function in $(0, \alpha]$. Moreover, we have

$$\begin{aligned} H'(\alpha) &= \frac{A(\alpha) + \alpha A'(\alpha)}{2} \\ &= \frac{1 - \phi^{-1}(\alpha)/\alpha + \phi^{-1}(\alpha)/\alpha - (\phi^{-1})'(\alpha)}{2} \\ &= \frac{1 - (\phi^{-1})'(\alpha)}{2} = \frac{1 - 1/\phi'(1/16)}{2}, \end{aligned}$$

and

$$\phi'(1/16) = \frac{4}{\pi} \int_0^{\frac{\pi}{2}} \frac{1}{1 + \sqrt{1 - \sin^2(t)}} = \frac{4}{\pi} \int_0^{\frac{\pi}{2}} \frac{1}{1 + \cos(t)}\, dt = \frac{4}{\pi}.$$

Hence, $H'(\alpha) = 1/2 - \pi/8 < 1$. Thus, $H'(\tau) < H'(\alpha) < 1$ contradicting the fact that $H'(\tau) = 1$. We conclude that the dominant singularity of $T$ appears when $T(x) = \alpha$. □

**Remark 4.3.6.** The value of $\rho$ is determined as follows: $T(\rho) = \alpha$, so since $T(x) = H(T(x)) + x + \frac{1}{2}T(x^2)$, we get $\rho + \frac{1}{2}T(\rho^2) = \alpha - H(\alpha)$, which can be solved numerically.

Finally, the following theorem gives an asymptotic expansion of $T(x)$ near $\rho$.

**Theorem 4.3.7.** *Let $\rho$ be as defined in Proposition 4.3.5. There exists $\theta'$ and $\epsilon'$ such that $T(x)$ is analytic in*

$$\Delta' = \{x \mid |x| < \rho + \epsilon' \text{ and } |Arg(x-\rho)| > \theta'\},$$

*and for $x \in \Delta'$, we have:*

$$T(x) = \alpha + C_1'(\rho - x) + C_2'(\rho - x)^2 + B'(\rho - x)^2 \log(\rho - x) + O(|(\rho - x)^3 \log(\rho - x)|). \quad (4.3.12)$$

*where $C_1', C_2'$ and $B'$ are constants that can be computed numerically, and $\alpha = \phi(1/16)$. Thus, the number of planar tanglegrams is asymptotically:*

$$[x^n]T(x) \sim C \cdot n^{-3} \cdot \rho^{-n}.$$

*Here, $\rho \approx 0.0634$, $\rho^{-1} \approx 15.7727$ and $C \approx 0.0078$ .*

*Proof.* Let $x \in B(0, \rho)$, and let $t_n$ be the $n^{th}$ coefficient of $T$. Since $T$ has positive coefficients, we have

$$|T(x)| = |\sum_{n \geq 0} t_n x^n| \leq \sum_{n \geq 0} t_n |x^n| < \sum_{n \geq 0} t_n \rho^n = \alpha.$$

Hence $T(x) \in B(0, \alpha)$. Moreover, by the implicit function theorem, $T$ can be continued analytically around each point $x$ of the circle $C(0, \rho)$ of center 0 and radius $\rho$, except perhaps around $\rho$. However around $\rho$, $T$ can be continued by Theorem 4.3.3. Thus, it is indeed possible to find $\epsilon'$ and $\theta'$ such that $T$ is analytic in $\Delta'$ as required. Now, by Theorem 4.3.3, we have

$$H(x) = C_0 + C_1(\alpha - x) + C_2(\alpha - x)^2 + B(\alpha - x)^2 \log(\alpha - x) + O(|(\alpha - x)^3 \log(\alpha - x)|).$$

By Theorem 4.2.3, the functional equation for $T(x)$ is given by

$$T(x) = H(T(x)) + x + \frac{T(x^2)}{2}.$$

Let $G(x) = x + \frac{T(x^2)}{2}$. By Remark 4.3.4, $G(x)$ is still analytic around $\rho$. Hence, the Taylor expansion of $G(x)$ gives

$$G(x) = G(\rho) + G'(\rho)(x - \rho) + \frac{G''(\rho)}{2}(x - \rho)^2 + O(|(x - \rho)^3|).$$

For simplicity, we let

- $D_0 = G(\rho) = \rho + \frac{T(\rho^2)}{2}$,
- $D_1 = -G'(\rho) = -(1 + \rho T'(\rho^2))$ and
- $D_2 = \frac{G''(\rho)}{2} = \frac{T'(\rho)}{2} + \rho^2 T''(\rho^2)$.

Since $\rho^2 < \rho$, $T'(\rho^2)$ and $T''(\rho^2)$ exist. Then $D_0, D_1$ and $D_2$ can be determined numerically. Thus,

$$\begin{aligned}T(x) &= C_0 + C_1(\alpha - T(x)) + C_2(\alpha - T(x))^2 \\ &+ B(\alpha - T(x))^2 \log(\alpha - T(x)) + O(|(\alpha - T(x))^3 \log(\alpha - T(x))|) \\ &+ D_0 + D_1(\rho - x) + D_2(\rho - x)^2 + O(|(\rho - x)^3|). \end{aligned} \tag{4.3.13}$$

The most important coefficient of the asymptotic expansion of $T(x)$ is the coefficient of

$$(\rho - x)^2 \log(\rho - x),$$

which will give us the asymptotic number of planar tanglegrams. We note that when $x \to \rho$ then $T(x) \to T(\alpha)$, hence $C_0 + D_0 = \alpha$. Again, as in the proof of Theorem 4.3.3, we will use bootstraping means to determine the coefficients. Some of the details are omitted since the computations are performed in the same way as in the proof of Theorem 4.3.3. First, we rewrite Equation (4.3.13):

$$T(x) = \alpha + C_1(\alpha - T(x)) + O(|(\alpha - T(x))^2 \log(\alpha - T(x))|) + D_1(\rho - x) + O(|(\rho - x)^2|).$$

Then,

$$T(x) = \alpha + \frac{D_1}{1 + C_1}(\rho - x) + O(|(\rho - x)^2 \log(\rho - x)|).$$

A second bootstrapping yields

$$T(x) = \alpha + \frac{D_1}{1 + C_1}(\rho - x) + \frac{B \cdot D_1^2}{(1 + C_1)^3}(\rho - x)^2 \log(\rho - x) + O(|(\rho - x)^2|)$$

and a third bootstrapping gives us Equation (4.3.12). Again as in the proof of Theorem 4.3.3, only $(\rho - x)^2 \log(\rho - x)$ contributes to $[x^n]H(x)$ when $n$ is large. Then, by [15, TheoremVI.2, Special cases] and [15, TheoremVI.9], we have

$$[x^n]T(x) = C \cdot \rho^{-n} \cdot n^{-3},$$

where

$$C = \frac{-2 \cdot B \cdot D_1^2 \cdot \rho^2}{(1 + C_1)^3}.$$

Here $C > 0$ because $B = -\frac{\pi^2}{16} < 0$ and $C_1 = \frac{\pi}{8} - \frac{1}{2}$. □

# List of References

[1] Mukul S Bansal, Wen-Chieh Chang, Oliver Eulenstein, and David Fernández-Baca. Generalized binary tanglegrams: Algorithms and applications. In *BICoB*, pages 114–125. Springer, 2009.

[2] Robert G Beiko, Timothy J Harlow, and Mark A Ragan. Highways of gene sharing in prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, 102(40):14332–14337, 2005.

[3] Sara C Billey, Matjaž Konvalinka, and Frederick A Matsen. On the enumeration of tanglegrams and tangled chains. *Journal of Combinatorial Theory, Series A*, 146:239–263, 2017.

[4] Sebastian Böcker, Falk Hüffner, Anke Truss, and Magnus Wahlström. A faster fixed-parameter approach to drawing binary tanglegrams. In *IWPEC*, pages 38–49. Springer, 2009.

[5] Miklós Bóna and Philippe Flajolet. Isomorphism and symmetries in random phylogenetic trees. *Journal of Applied Probability*, 46(4):1005–1019, 2009.

[6] John Adrian Bondy and Uppaluri Siva Ramachandra Murty. *Graph theory with applications*. Citeseer, 1976.

[7] Kevin Buchin, Maike Buchin, Jaroslaw Byrka, Martin Nöllenburg, Yoshio Okamoto, Rodrigo I Silveira, and Alexander Wolff. Drawing (complete) binary tanglegrams. *Algorithmica*, 62(1):309–332, 2012.

[8] Reinhard Diestel. *Graph theory*. Springer-Verlag Berlin and Heidelberg GmbH & amp, 2000.

[9] NIST Digital Library of Mathematical Functions. http://dlmf.nist.gov/, Release 1.0.15 of 2017-06-01. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[10] Michael Drmota. *Random trees: an interplay between combinatorics and probability*. Springer Science & Business Media, 2009.

[11] David Steven Dummit and Richard M Foote. *Abstract algebra*, volume 3. Wiley Hoboken, 2004.

[12] Henning Fernau, Michael Kaufmann, and Mathias Poths. Comparing trees via crossing minimization. *FSTTCS 2005: Foundations of Software Technology and Theoretical Computer Science*, pages 457–469, 2005.

[13] Philippe Flajolet, Zhicheng Gao, Andrew Odlyzko, and Bruce Richmond. The distribution of heights of binary trees and other simple trees. *Combinatorics, Probability and Computing*, 2:145–156, 1993.

[14] Philippe Flajolet and Andrew Odlyzko. The average height of binary trees and other simple trees. *Journal of Computer and System Sciences*, 25(2):171–213, 1982.

[15] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University press, 2009.

[16] Marshall Hall. The theory of groups. *The Macmillan Company*, 1959.

[17] F. Harary. *Graph Theory*. Addison-Wesley series in mathematics. Avalon Publishing, 1969.

[18] Frank Harary and Edgar M. Palmer. *Graphical Enumeration*. Academic Press, 1973.

[19] Israel N Herstein. *Topics in algebra*. John Wiley & Sons, 1964.

[20] Camille Jordan. Sur les assemblages de lignes. *Journal für die reine und angewandte Mathematik*, 70:185–190, 1869.

[21] Matjaž Konvalinka and Stephan Wagner. The shape of random tanglegrams. *Advances in Applied Mathematics*, 78:76–93, 2016.

[22] Antoni Lozano, Ron Y Pinter, Oleg Rokhlenko, Gabriel Valiente, and Michal Ziv-Ukelson. Seeded tree alignment and planar tanglegram layout. In *International Workshop on Algorithms in Bioinformatics*, pages 98–110. Springer, 2007.

[23] Frederick Matsen, Sara Billey, Arnold Kas, and Matjaž Konvalinka. Tanglegrams: a reduction tool for mathematical phylogenetics. *IEEE/ACM transactions on computational biology and bioinformatics*, 2016.

[24] Frank WJ Olver, Daniel W Lozier, Ronald F Boisvert, and Charles W Clark. *NIST handbook of mathematical functions*. Cambridge University Press, 2010.

[25] Roderic DM Page. Parasites, phylogeny and cospeciation. *International Journal for Parasitology*, 23(4):499–506, 1993.

[26] Roderic DM Page. *Tangled trees: phylogeny, cospeciation, and coevolution*. University of Chicago Press, 2003.

[27] Helmut Prodinger. A note on the distribution of the three types of nodes in uniform binary trees. *Séminaire Lotharingien de Combinatoire 38b*, 1996.

[28] Dimbinaina Ralaivaosaona, Jean Bernoulli Ravelomanana, and Stephan Wagner. Enumeration of planar tanglegrams. In preparation.

[29] Celine Scornavacca, Franziska Zickmann, and Daniel H Huson. Tanglegrams for rooted phylogenetic trees and networks. *Bioinformatics*, 27(13):i248–i256, 2011.

[30] Neil JA Sloane and James A Sellers. On non-squashing partitions. *Discrete mathematics*, 294(3):259–274, 2005.

[31] N.J.A Sloane and editor. The on-line encyclopedia of integer sequences, 2017.

[32] W. A. Stein et al. *Sage Mathematics Software (Version 7.2.0)*. The Sage Development Team, 2016.

[33] Balaji Venkatachalam, Jim Apple, Katherine St John, and Daniel Gusfield. Untangling tanglegrams: comparing trees by their drawings. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 7(4):588–597, 2010.

[34] Christopher Whidden, Norbert Zeh, and Robert G Beiko. Supertrees based on the subtree prune-and-regraft distance. *Systematic biology*, 63(4):566–581, 2014.

[35] Hassler Whitney. Non-separable and planar graphs. *Transactions of the American Mathematical Society*, 34, 04 1932.